IMAGE ANALYSIS AND MODELING

Paul A. Wintz, et al

Purdue University

Prepared for:

Rome Air Development Center
Defense Advanced Research Projects Agency

August 1975

RADC-TR-75-202
Quarterly Technical Report
August 1975

IMAGE ANALYSIS AND MODELING

Purdue University

Sponsored by
Defense Advanced Research Projects Agency
ARPA Order No. 2893

Approved for public release;
distribution unlimited.

The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the Defense Advanced Research Projects Agency or the U. S. Government.

Rome Air Development Center
Air Force Systems Command
Griffiss Air Force Base, New York 13441

This report has been reviewed by the RADC Information Office (OI) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public including foreign nations.

This technical report has been reviewed and is approved for publication.

APPROVED: John M. Trossbach, Jr

JOHN M. TROSSBACH, Jr., Capt, USAF
Project Engineer

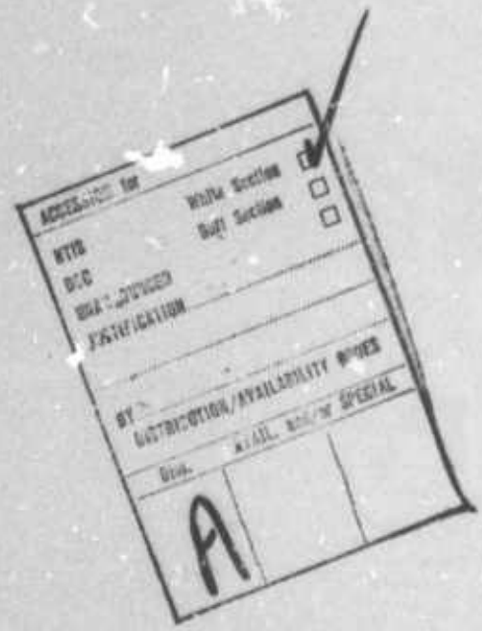


IMAGE ANALYSIS AND MODELING

Paul A. Wintz
T. S. Huang

Contractor: Purdue University
Contract Number: F30602-75-C-0150
Effective Date of Contract: 27 February 1975
Contract Expiration Date: 31 October 1975
Amount of Contract: $427,020.00
Program Code Number: 5D30
Period of work covered: 1 February 1975 - 30 April 1975

Principal Investigator: Paul A. Wintz and T. S. Huang
Phone: 317 493-3361

Project Engineer: Capt John M. Trossbach, Jr., USAF
Phone: 315 330-3175

Approved for public release;
distribution unlimited.

This research was supported by the Defense Advanced Research Projects Agency of the Department of Defense and was monitored by Capt John M. Trossbach, Jr., RADC (IRRO), Griffiss AFB NY 13441.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER RADC-TR-75-202	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) IMAGE ANALYSIS AND MODELING		5. TYPE OF REPORT & PERIOD COVERED Quarterly Technical Report 1 Feb 75 - 30 Apr 75
		6. PERFORMING ORG. REPORT NUMBER N/A
7. AUTHOR(s) Paul A. Wintz T. S. Huang		8. CONTRACT OR GRANT NUMBER(s) F30602-75-C-0150
9. PERFORMING ORGANIZATION NAME AND ADDRESS Purdue University School of Electrical Engineering West Lafayette IN 47907		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61101E 28930001
11. CONTROLLING OFFICE NAME AND ADDRESS Defense Advanced Research Projects Agency 1400 Wilson Blvd Arlington VA 22209		12. REPORT DATE August 1975 13. NUMBER OF PAGES ~~190~~ 195
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) Rome Air Development Center (IRRO) Griffiss AFB NY 13441		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE N/A

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

Same

18. SUPPLEMENTARY NOTES

RADC Project Engineer:
John M. Trossbach, Jr., Capt, USAF (IRRO)

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Image Processing, Digital Image Processing, Image Restoration, Image Coding, Pattern Recognition, Syntactic Pattern Recognition, Image Segmentation

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report summarizes the results of the research program on Image Analysis and Modeling supported by the Advanced Research Projects Agency under Contract No. F30602-75-C-0150 for the three month period February 1 to April 30, 1975. The program started October 1, 1973. The objective is to achieve a better understanding of image structure and to use this knowledge to develop improved models for use in image analysis and processing tasks such as information extraction, image enhancement and restoration, coding, etc. The ultimate objective of this research is to form a basis for the development of technology

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

relevant to military applications of machine extraction of information from aircraft and satellite imagery of the earth's surface.

TABLE OF CONTENTS

IMAGE ANALYSIS AND MODELING

Research Summary and Overview

In this report we summarize our research activities during the three month period February 1 - April 30, 1975. Our research program started October 1, 1973. During the first year we concentrated our efforts on basic research in image structure analysis and image modeling. During the second year we are continuing our basic research, but in the meantime also look into several applications in the information extraction area. We hope that on the one hand the results of our basic research could find practical applications, and on the other hand problems arising from practical situations would guide the direction of our basic research.

In basic research, our emphasis has been shifted toward image understanding and information extraction. The ultimate goal in automatic image understanding and information extraction is to build a machine which will look at a scene and then give us a qualitative as well as quantitative description of it. This formidable task can be broken down into several sub-tasks as indicated by the block diagram in Fig. 1.

The sensor looks at the scene and outputs image data which are generally multispectral and multitemporal. The preprocessing box attempts either to put the image into a form which is more suitable for analysis by the latter boxes or to compress the image data for transmission to a remote location or storage for future use. The image segmentation box, in simple instances, locates objects in the image (for example, in character recognition, it locates the characters). For more complex scenes, it segments the image into regions. Each of these regions is classified by the classification box. The classification can be done either by classical decision-theoretic methods (feature extraction followed by statistical classification) or by the more recent

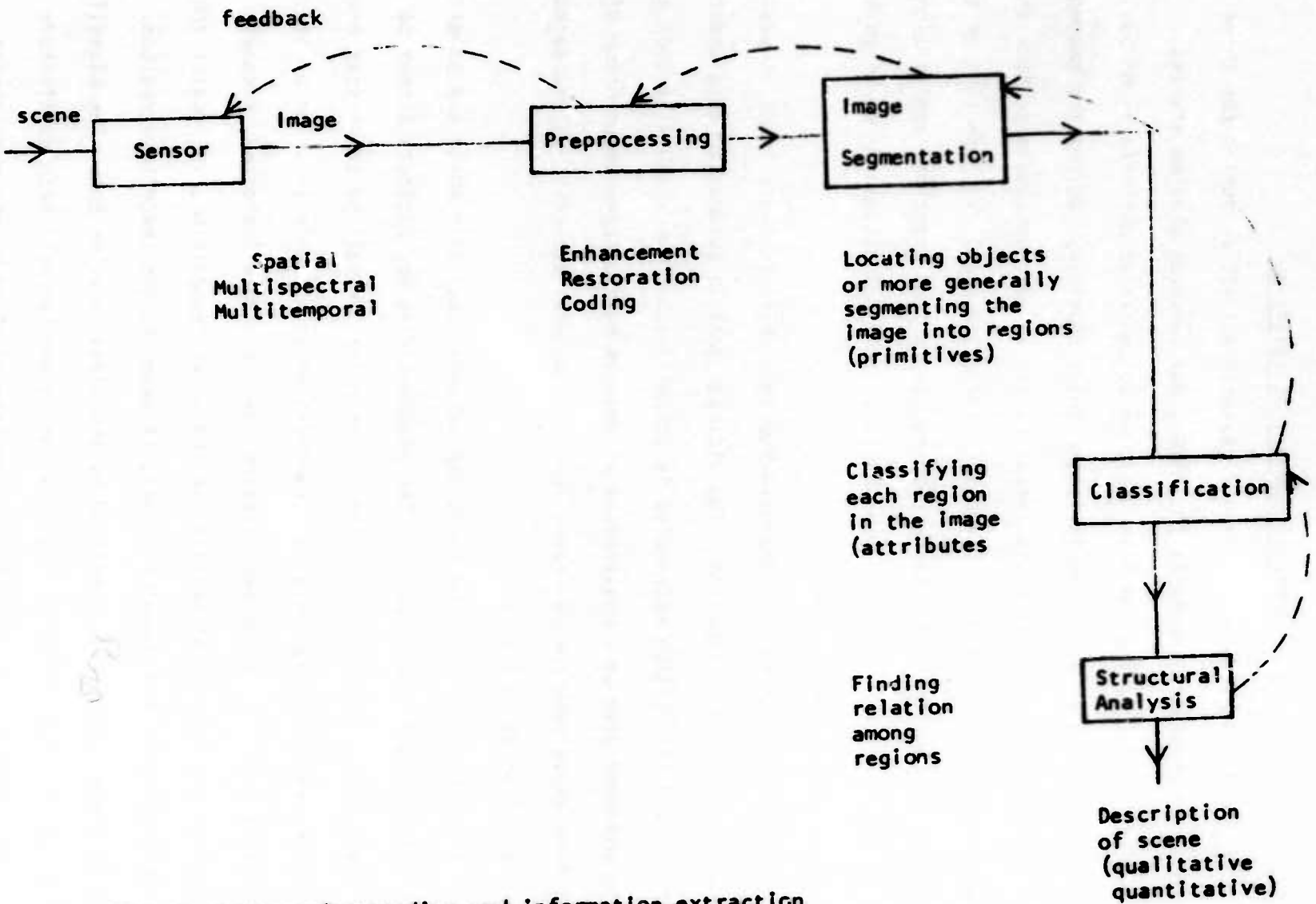


Fig. 1 Image understanding and information extraction

syntactical methods. In linguistic terminology, the regions are primitives, and the classifier finds attributes for these primitives. Finally, the structural analysis box attempts to find the relationship (spatial, temporal, spectral) among the regions (primitives). In some cases, one may want to develop a grammar for the primitives.

The functions of the boxes in Fig. 1 overlap, and their demarcation is not clear cut. Furthermore, there is strong feedback in the system (as indicated by the dashed arrows in the block diagram). For example, structural analysis may reveal that certain regions are classified wrongly and thus force a reclassification. Similarly, difficulties encountered in classification may make it necessary to reexamine image segmentation.

The research projects we are carrying and will carry out fall into six overlapping categories corresponding to the five boxes in Fig. 1 and applications. In the present report our emphasis is on applications, image segmentation, and pattern classification. We mention some highlights below.

We have attacked quite successfully four application-oriented problems: locating large airports in ERTS imagery (Huang and Chan), detecting water boundaries in ERTS imagery for flood maps (Chen and Wintz), change detection in missile movie frames (Huang and Burnett), and TV bandwidth compression for remotely piloted vehicle images (Wintz).

In image segmentation we are pursuing two approaches: edge detection, and region analysis. In the case of edge detection we are particularly interested in developing techniques which would work even if the image is noisy and smeared (Yang and Huang, Burnett and Huang). The investigation on the use of syntactic approach (Kelly, in the last Quarterly Progress Report) is being continued.

In the case of region analysis we have developed earlier the BLOB algorithm (Wintz and Gupta, Semi-annual Report, Oct. 1, 1973 - March 31, 1974). We are currently working on improved versions of BLOB, and in the meantime looking into techniques which would give results that are independent of the order of processing. One such technique is the valley-seeking unsupervised clustering method of Fukunaga and Narendra developed recently at Purdue. The use of unsupervised clustering methods for image segmentation looks promising (Yoo and Huang). The success of region analysis depends critically on a suitable choice of feature measurements. The most important class of features for this application is perhaps texture descriptors. We have been studying texture descriptors intensively and have developed several new ones which give good classification results (Mitchell and Myers).

In pattern classification one is often overwhelmed by a large number of potentially useful features. On the other hand, one can usually afford to use only a few. The key question is how to select m features from a potential set of n (>m) in an optimal way. We have developed a very efficient method of feature selection which is orders of magnitude faster than exhaustive search (Fukunaga and Narendra).

Most of the research projects reported in the following pages are continuing.

LOCATE AIRPORTS IN ERTS IMAGERY

T. S. Huang and C. K. Chan

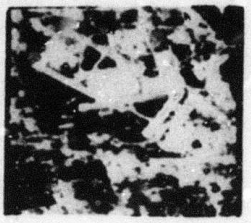
I. OBJECTIVE

The goal of this project is to develop computer techniques for locating large airports in ERTS multispectral imagery.

II. PROGRESS:

Several large airports in the United States (O'hare Airport in Chicago, San Francisco Airport, etc.) are picked out from ERTS data to provide a preliminary study of their texture and properties. It is found that in all four spectral channels used by ERTS, the spectral reflectivity of the airports and highways are very similar. Thus a preliminary objective is to pick out and separate the airports and highways (i.e., concrete) from the other contents of the ERTS data. A statistical pattern recognition approach is used in this study making use of the facilities in LARS. A training set of data for concrete is hand-picked and the rest of the picture is classified as concrete or not-concrete from the statistics of the training set. Fig. 1 shows the picture of O'hare Airport in channel 2 (wavelength 0.69 μm) and Fig. 2 shows the classified and thresholded four channels output. A threshold of ℓ means that all points having a probability of less than $\ell/100$ of being concrete is thresholded away (black part of picture) and points with probability $> \ell/100$ of being concrete shows up in white in the picture.

Since airports and highways usually are composed of straight lines, a Fourier transform approach is used to process the classified data. A Fourier transform of Fig. 2b is shown in Fig. 3. The Fourier transform is then filtered so that only the wedge corresponding to the parallel runways are left.



O'hare Airport

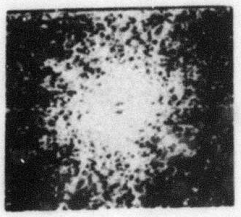
Fig. 1

ℓ=5

Fig. 2a

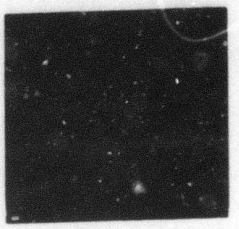
ℓ=10

Fig. 2b



Fourier transform of Fig. 2b

Fig. 3



Thresholded inverse transform of Fig. 3

Fig. 4

Thresholded single channel (Ch.2) output

Fig. 5

Inverse transform is then performed and the results thresholded. See Fig. 4. This study provides an idea of how well one can identify airport runways from a picture like Fig. 1.

The Fourier transform approach is also applied to the data in a single channel (ch. 2). The thresholded result is shown in Fig. 5. Although it came out better than the thresholded 4 channels classification results, we feel that a better choice of training set would greatly improve the result in Fig. 4.

III. PROPOSAL FOR FURTHER WORK

So far, the study has emphasized removing unwanted information from a picture so as to bring out the runways. The filter used for the Fourier transform and the threshold used on the resulting picture are handpicked. Further work will be aimed at automating these choices.

The first thing that will be done is to threshold the Fourier transform in Fig. 3 so that only the prominent lines are left. The inverse transform can then be taken so that hopefully only the straight lines that signify the highways and runways will remain. An automatic thresholding scheme can then be used to produce results similar to that of Fig. 5. If this proves successful, larger sections of different airports will be tried.

BOUNDARY DETECTION FOR U.S. ARMY CORPS OF ENGINEERS FLOOD MAPS

P. H. Chen and P. A. Wintz

The U.S. Army Corps of Engineers (USACE) is in charge of flood control in the United States. In order to determine which flood gates to open or close and in order to predict river flood conditions they use mathematical models. The inputs to the mathematical models are measurements taken from river water level water gauges, rain gauges, snow gauges, etc. There is reason to believe that the data required by their models could be obtained in a more cost effective manner from satellite pictures. For example, precipitation maps and maps showing changes in snow cover are indicative of the amount of water inputted to the water shed. Furthermore, flood area extent maps provide required data.

The purpose of this project is to determine flood area extent maps from ERTS multispectral scanner data and to code these maps in an efficient manner in order to reduce the bulk of the data inputted to USACE mathematical models. A few preliminary results are illustrated in Figs. 1-5. Fig. 1 shows band 7 of an ERTS picture encompassing part of the Ohio River. Water shows up dark in band 7 since infrared energy is absorbed by water. Maximum likelihood classifier was trained on some river and non-river data and then used to classify each picture element into either water or not water. The resulting classification map is presented in Fig. 2 where water was made bright and non-water picture elements dark. Fig. 3 illustrates a classification map of the same area when there is flood. Figs. 4 and 5 illustrate the classification results of another section of the river taken at two different times (no flooding and flooding). Preliminary results indicate that 2-dimensional runlength codes and contour codes will significantly reduce the amount of data

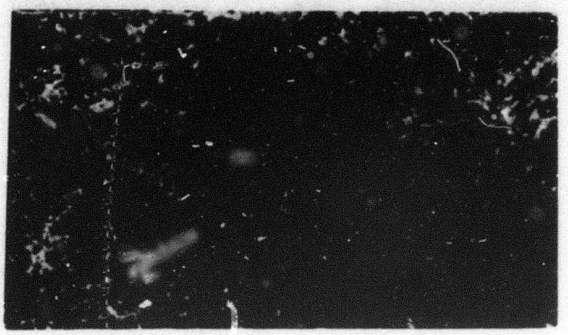


Fig. 1 ERTS Run 72062801
C(2454, 3098, 1)

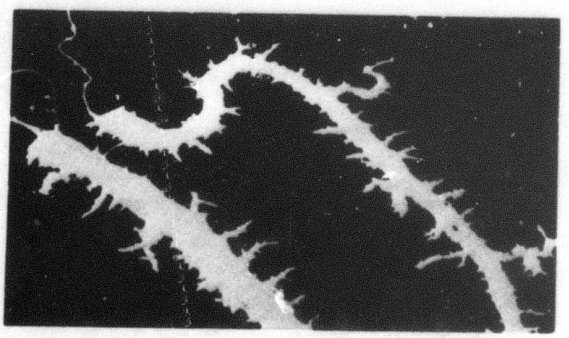


Fig. 2 Classification Result
of Fig. 1

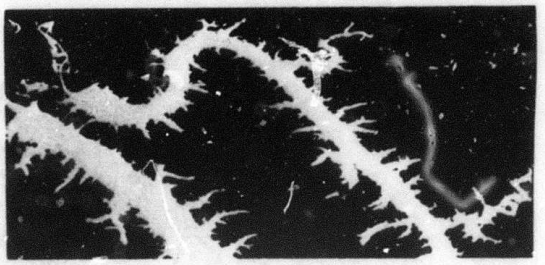


Fig. 3 Classification Result (Flooding)
of the same area as Fig. 1

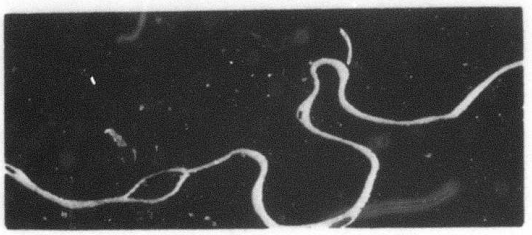


Fig. 4 Classification Result (No Flood)
of one section of the river

Fig. 5 Classification Result (Flooding)
of the same area as Fig. 4

that must be processed by the USACE mathematical models. This is extremely important since the Mississippi Valley watershed covers about 129 ERTS frames consisting of approximately 30 million bits.

REFERENCE

[1] T. S. Huang, W. F. Schreiber, and O. J. Tretiak, "Contour Coding of Images," Picture Bandwidth Compression, ed. by T. S. Huang and O. J. Tretiak, pp. 443-448, 1972.

EXPERIMENTS IN CHANGE DETECTION IN MISSILE IMAGERY

T. S. Huang and J. W. Burnett

I. INTRODUCTION

In Figures 1(a) and (b) we have two successive frames (digitized) of a movie film of a missile. The original film was supplied by the U. S. Army White Sand Missile Range. We report here some experiments in automatic detection of change between the two frames. Change detection has obvious applications in the tracking of moving objects.

II. CORRELATION AND DIFFERENCING

Each frame was digitized to 128x128 samples, and 8 bits/sample. The two frames were lined up by cross-correlation. Then, in the first experiment, one frame was subtracted from the other point by point, resulting in Figure 2. Thresholding Figure 2 yielded Figure 3. The missile motion was detected, but the result contained a considerable amount of extraneous points.

III. EDGE COMPARISON

In the second experiment, we first extract the edge points (using a smoothed gradient operator) in the two frames, yielding Figures 4(a) and (b). Then an "exclusive-or" operation was performed (point by point) on these two edge pictures, resulting in Figure 5. A point in Figure 5 was white if and only if the corresponding points in Figs. 4(a) and (b) were of different colors (one black and the other white). Again the missile motion was detected, but again we have a large amount of extraneous points.

IV. ELASTIC EDGE COMPARISON

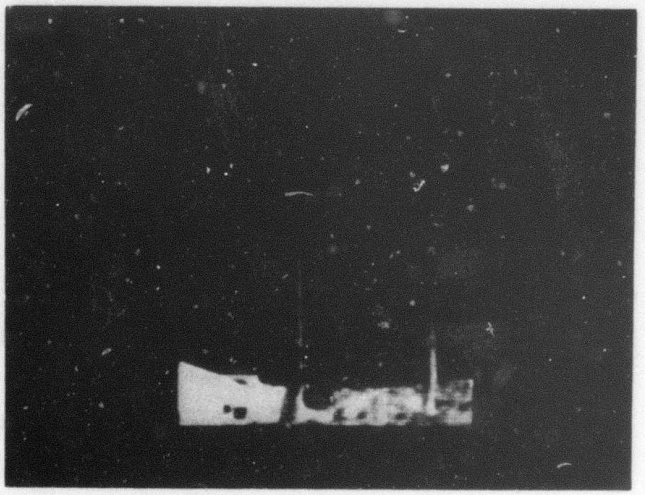
To reduce extraneous points in the result, we performed a third experiment in which the two edge pictures (Figures 4(a) and (b)) were compared with some tolerance allowed. Specifically, for each black point in Figure 4(b), we examined the 3 x 3 point neighborhood around the corresponding point in

Figure 4(a). We painted the corresponding point in Figure 6 black if and only if the 3 x 3 neighborhood contained no black points. In the resulting Figure 6 we observe that the missile motion was detected, and the extraneous points appeared in small groups (3 points or less in each connected group). These extraneous points can be eliminated completely if we erase all connected point groups which contain 3 or less points.

V. CONCLUSION

The elastic edge comparison technique seems to work well in detecting changes between two scenes which are essentially identical except for the changing parts.

(a)



(b)

Fig. 1 Two successive frames (digitized) of a missle movie film.

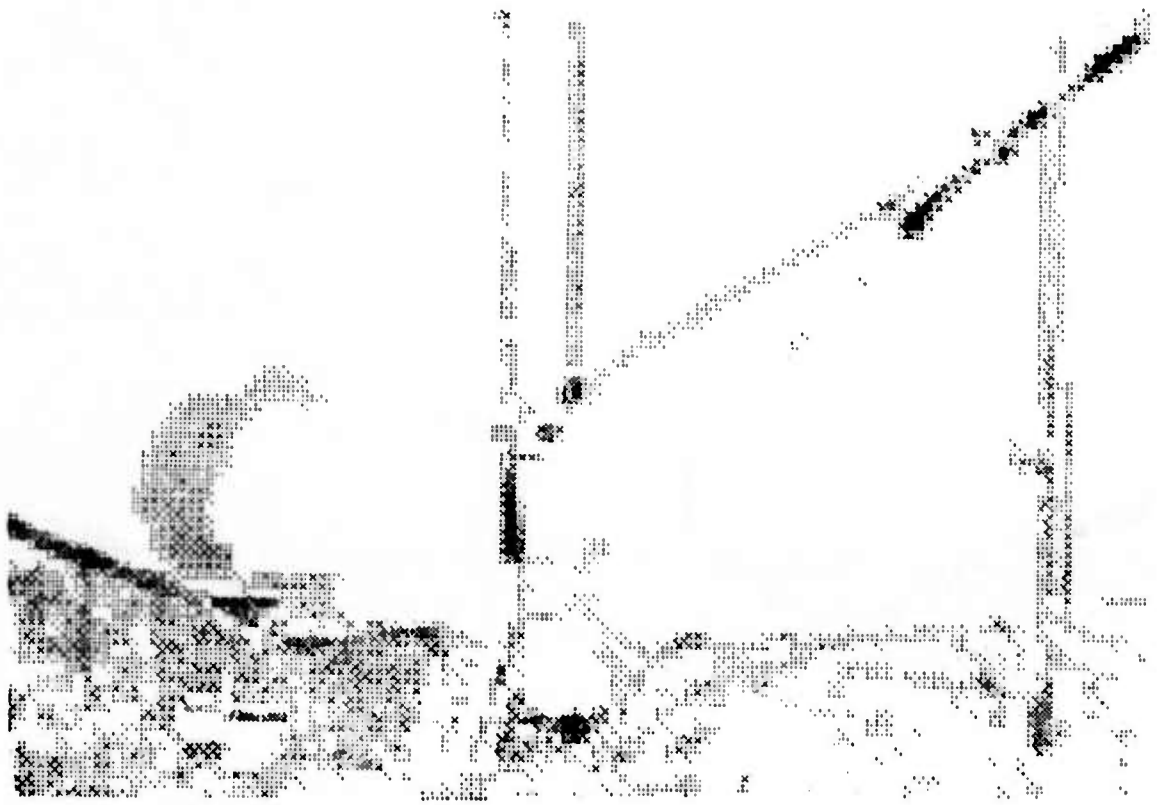


Fig. 2 Difference between two frames

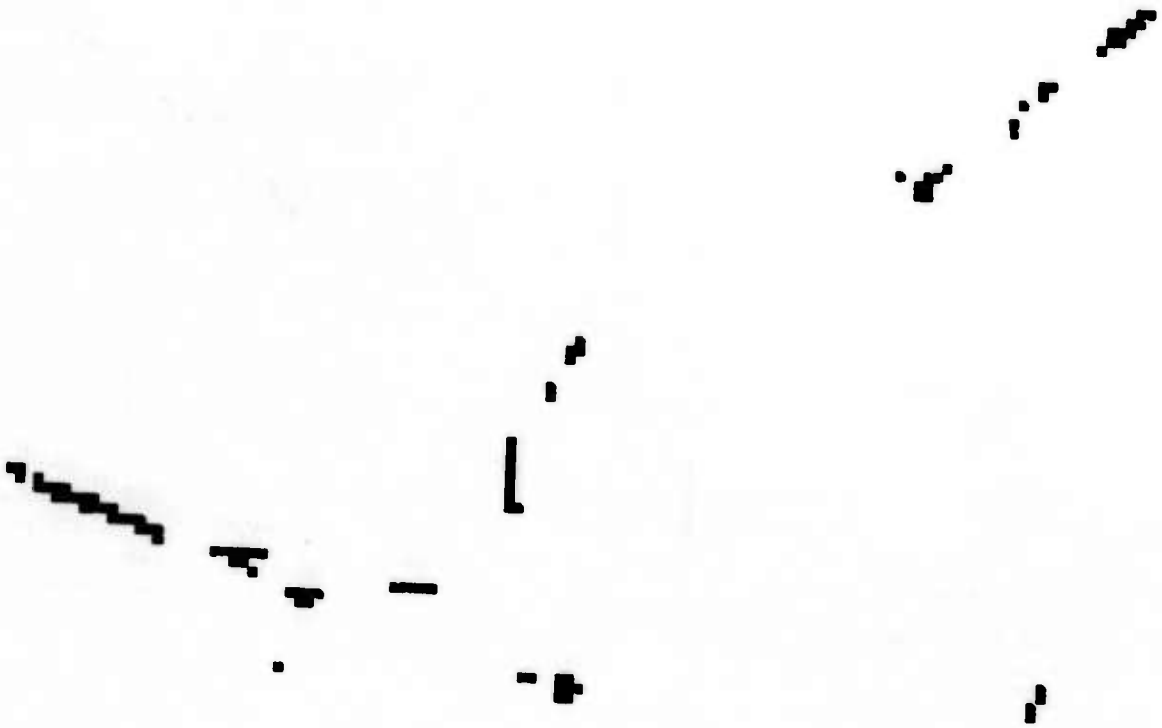
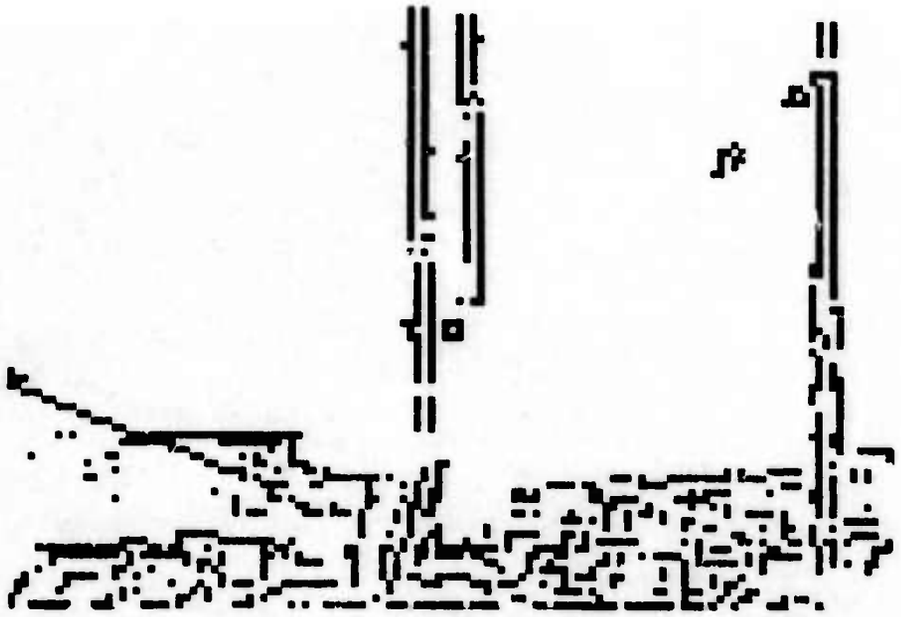
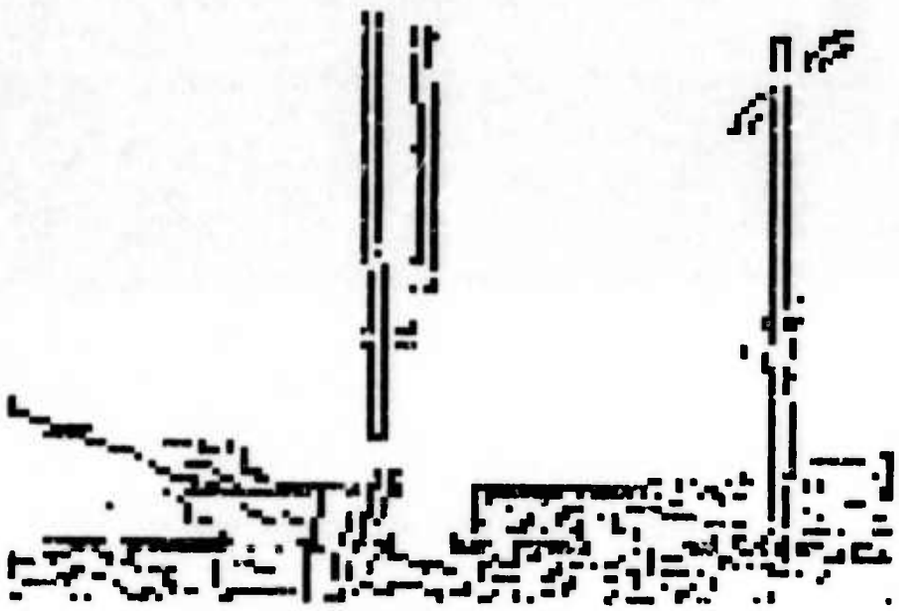


Fig. 3 Thresholded difference



(a)



(b)

Fig. 4 Edge pictures

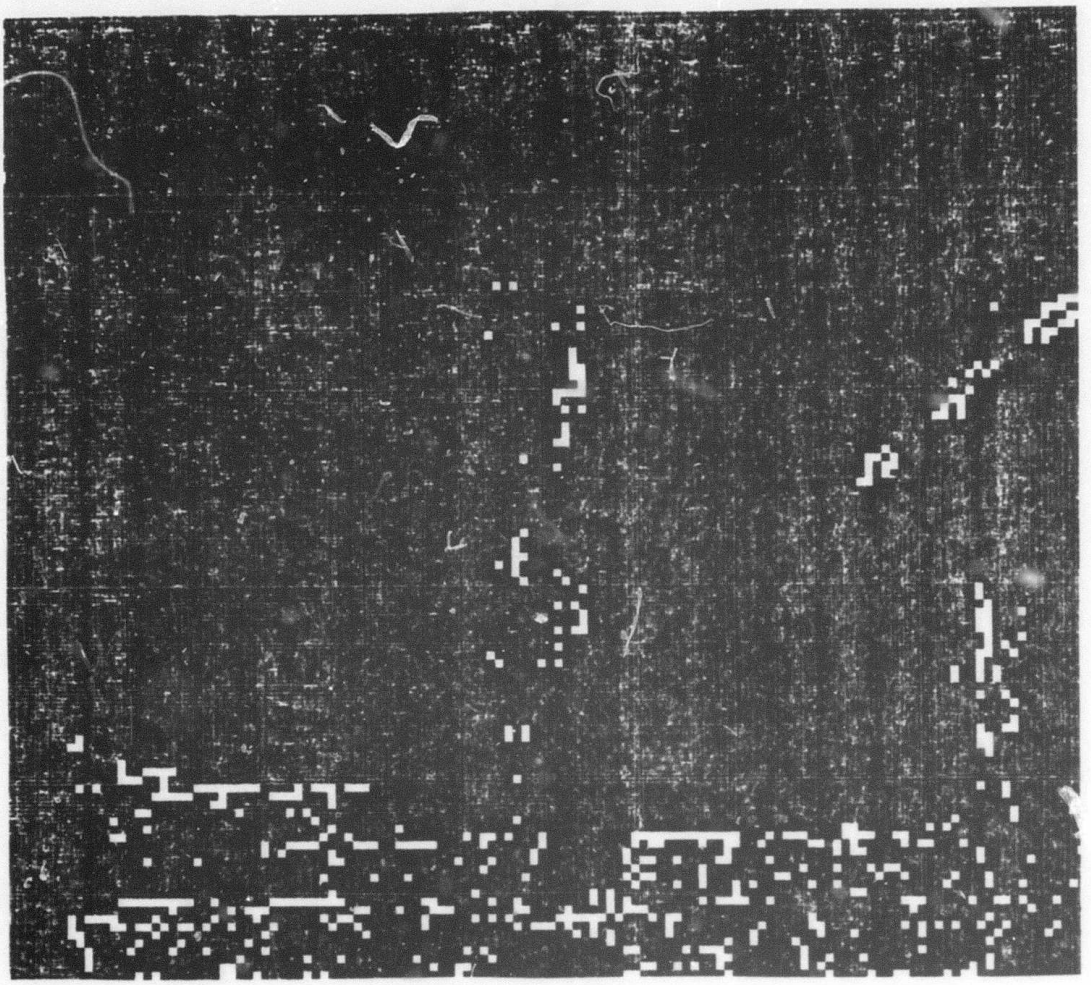


Fig. 5 Result of edge comparison

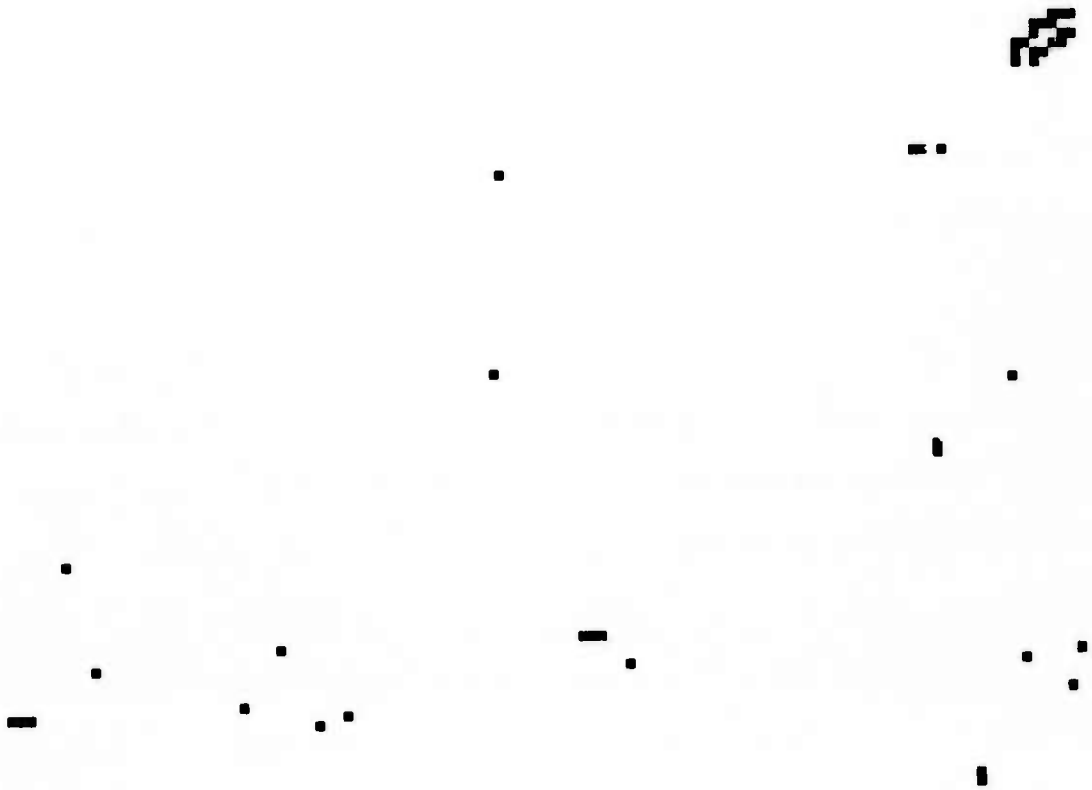


Fig. 6 Result of elastic edge comparison

TV BANDWIDTH COMPRESSION FOR RPV's

Paul A. Wintz

A critical problem in the RPV system is the bandwidth required to transmit the TV signals from the vehicle to the vehicle controller. There are three possibilities for reducing the bandwidth: (1) reduce the temporal resolution by transmitting less than 30 frames per second; (2) reduce the spatial resolution by coding fewer than all of the picture elements in the frames transmitted; (3) reduce the spatial redundancy in the remaining picture elements.

The purpose of this project is to investigate the tradeoffs between the latter two methods for reducing the number of bits required to code the picture.

The original picture was digitized into a 512x512 array of picture elements of 8 bits each. The resolution was then reduced to 256x256 picture elements by averaging subsets of 2x2 picture elements in the 512x512 picture. The 512x512 picture is presented in Fig. 1 and the 256x256 picture is presented in Fig. 5.

Each of these pictures was processed by the hybrid coding scheme due to Habibi. Each line of picture elements was first divided into sub-blocks containing 32 picture elements each. Each of these 1x32 arrays of picture elements was transformed with the cosine transformation to obtain 32 coefficients. This cosine transformation was performed on each 1x32 subsets of the entire original picture in order to reduce the redundancy in the horizontal direction. Successive scan lines are also highly correlated. Since the picture elements in a 1x32 sub-array of picture elements are highly correlated with the picture elements in the 1x32 sub-array on the next scan line, and since we do the same transformation on both sub-arrays, we also

expect the coefficients produced by the transformations to be highly correlated. This redundancy can be reduced by differencing the coefficients. This can be accomplished by the coding scheme known as DPCM. In summary, the cosine transform in the horizontal direction followed by DPCM in the vertical direction to obtain 1x32 arrays of the coefficient differences.

Since the coefficients have different variances it follows that the differences between coefficients will also have different variances. In order to efficiently code the coefficient differences we use block quantization. That is, we use a different number of bits to code the different coefficient differences. The bit assignment is optimized in order to minimize the RMS quantization error. The bit allocations are listed in Table 1.

The picture in Fig. 1 was coded by this method and then decoded to form a 512x512 replica of the original image. The decoded pictures for 2, 1, and 1/2 bits per picture element are presented in Figs. 2,3, and 4. The 2-bit picture has only a very small degradation relative to the original. The one-bit picture has more degradation. The 1/2-bit picture not only has very noticeable degradation but the edges of the 1x32 blocks are also apparent.

The same coding scheme was applied to the 256x256 picture of Fig. 5. The coded pictures were decoded to form a replica of the 256x256 image which was then converted to a 512x512 image by repeating each picture element. The results are shown in Figs. 6, 7, and 8 for 2, 1, and 1/2 bits per picture element. The same general comments apply as for the 512x512 picture except that the distortions are significantly more severe. The reason for this is that there is less correlation between the 256x256 picture elements than between the 512x512 picture elements and so the coding algorithm is expected to be less efficient.

Finally, we compare the 512x512 picture coded at 1/2-bit per picture

Coefficients	2 bits		1 bit		1/2 bit	
	512 picture	256 picture	512 picture	256 picture	512 picture	256 picture
1	4	4	3	3	3	3
2	4	3	3	2	2	2
3	3	3	2	2	2	2
4	3	3	2	2	1	1
5	3	3	2	2	1	1
6	3	3	2	2	1	1
7	3	3	2	2	1	1
8	3	3	2	2	1	1
9	3	3	2	2	1	1
10	3	2	2	1	1	1
11	3	2	2	1	1	1
12	2	2	1	1	1	1
13	2	2	1	1	0	0
14	2	2	1	1	0	0
15	2	2	1	1	0	0
16	2	2	1	1	0	0
17	2	2	1	1	0	0
18	2	2	1	1	0	0
19	2	2	1	1	0	0
20	1	2	0	1	0	0
21	1	2	0	1	0	0
22	1	2	0	1	0	0
23	1	1	0	0	0	0
24	1	1	0	0	0	0
25	1	1	0	0	0	0
26	1	1	0	0	0	0
27	1	1	0	0	0	0
28	1	1	0	0	0	0
29	1	1	0	0	0	0
30	1	1	0	0	0	0
31	1	1	0	0	0	0
32	1	1	0	0	0	0

Table 1 Bit allocation for coefficient differences

element illustrated in Fig. 4 with the 256x256 picture coded at 2 bits per picture element shown in Fig. 8. Both of these pictures require the same number of bits to transmit. The 512x512 picture at 1/2 bit is subjectively better than the 256x256 picture at 2 bits.

The same experiment was performed on a different picture and the decoded pictures are presented in Figs. 9 to 16. Fig. 9 is the 512x512 original, and Figs. 10, 11, and 12 are coded at 2 bits, 1 bit, and 1/2 bit respectively. Fig. 13 is the 256x256 original and Figs. 14, 15, and 16 correspond to coding this picture with 2 bits, 1 bit, and 1/2 bit. Careful observation of these pictures leads to the same conclusion as for the previous pictures.

Reference

A. Habibi, "Hybrid Coding of Pictorial Data," IEEE Trans. on Communications, Vol. COM-22, No. 5, pp. 614-624, May 1974.

Fig. 1 512x512 Original Image

Fig. 2 512x512 Coded Image, 2 bits/picture element

Fig. 3 512x512 Coded Image, 1 bit/picture element

Fig. 4 512x512 Coded Image, 1/2 bit/picture element

Fig. 5 256x256 Original Image

Fig. 6 256x256 Coded Image, 2 bits/picture element

Fig. 7 256x256 Coded Image, 1 bit/picture element

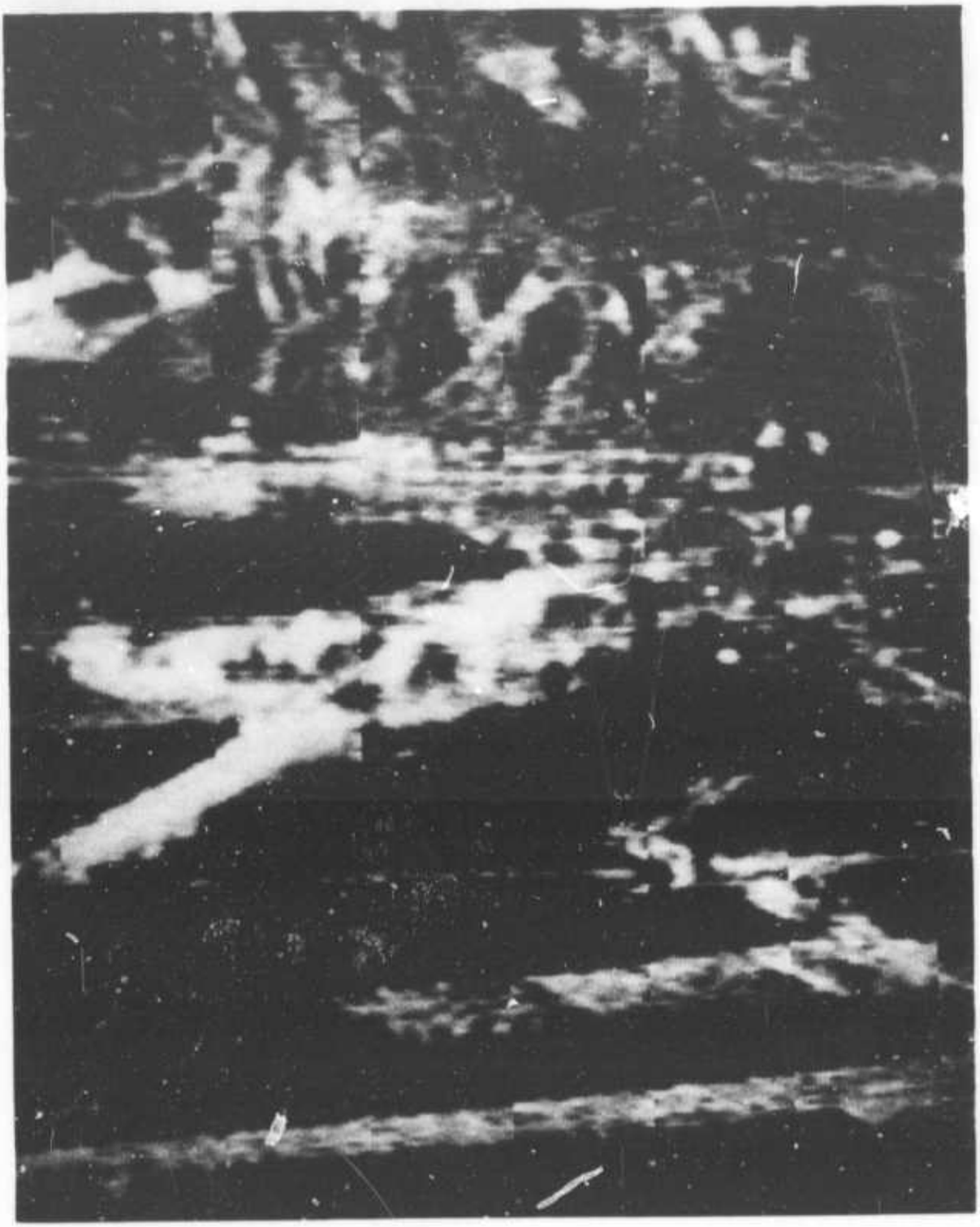


Fig. 8 256x256 Coded Image, 1/2 bit/picture element

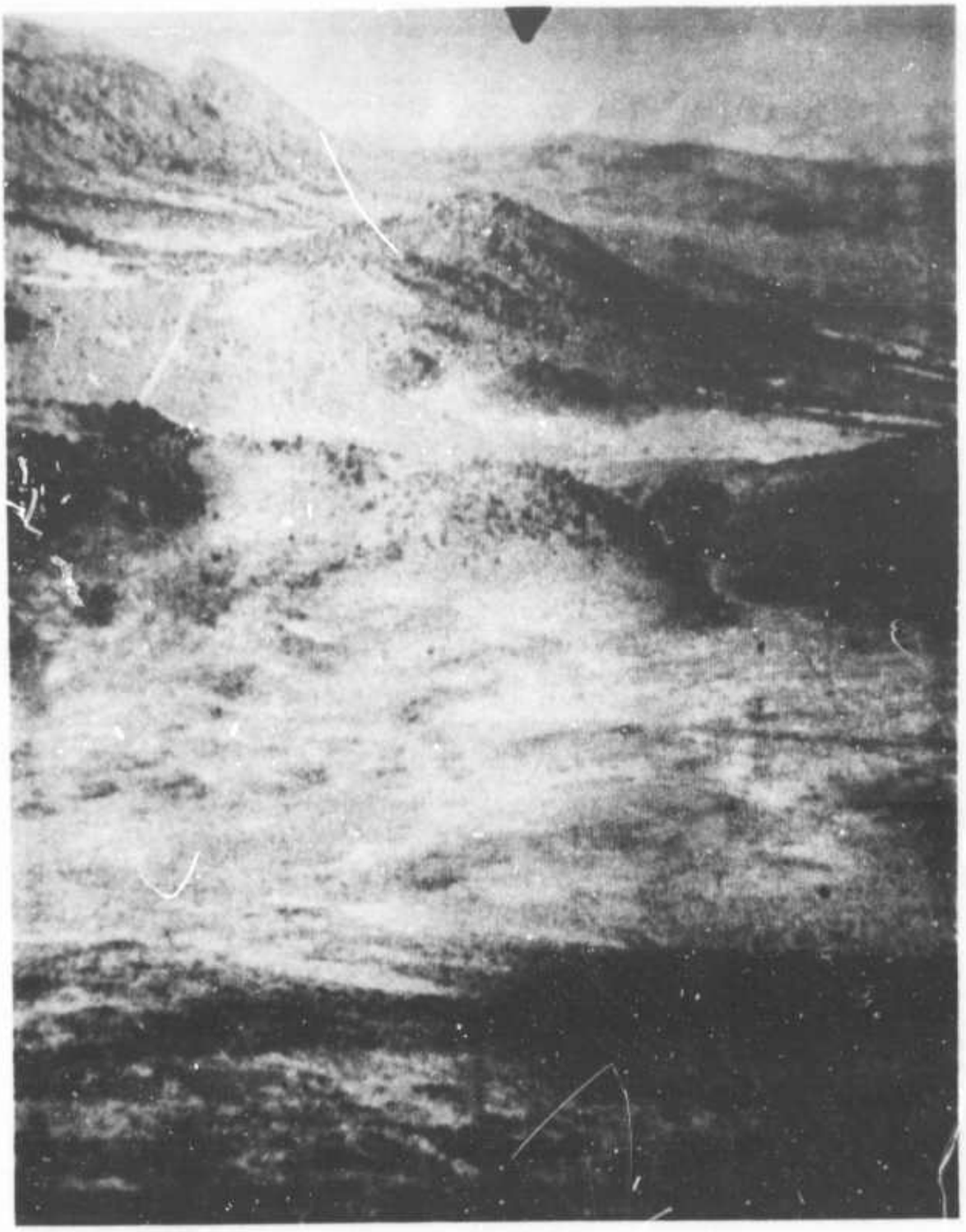


Fig. 9 512x512 Original Image

Fig. 10 512x512 Coded Image, 2 bits/picture element

Fig. 11 512x512 Coded Image, 1 bit/picture element

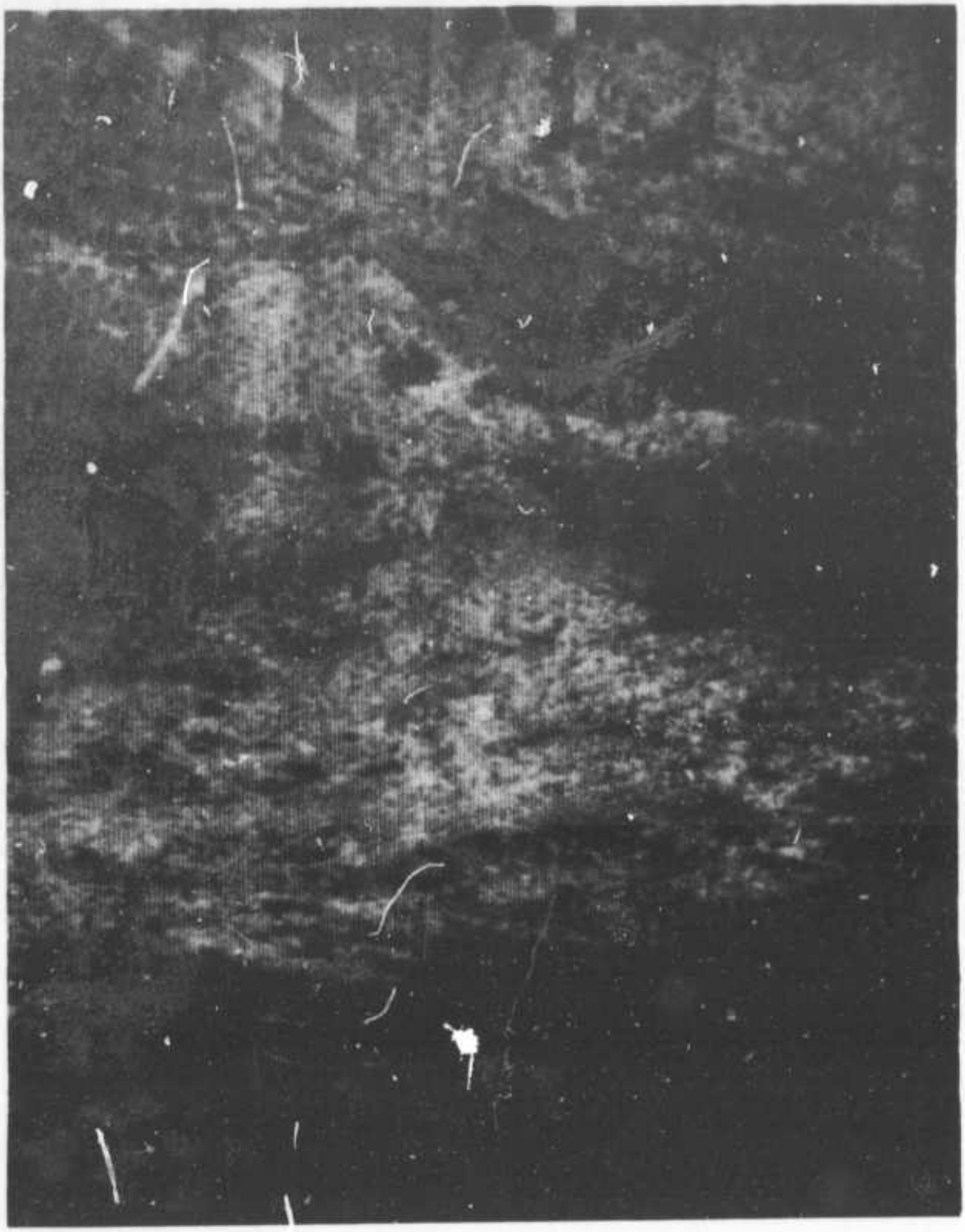


Fig. 12 512x512 Coded Image, 1/2 bit/picture element

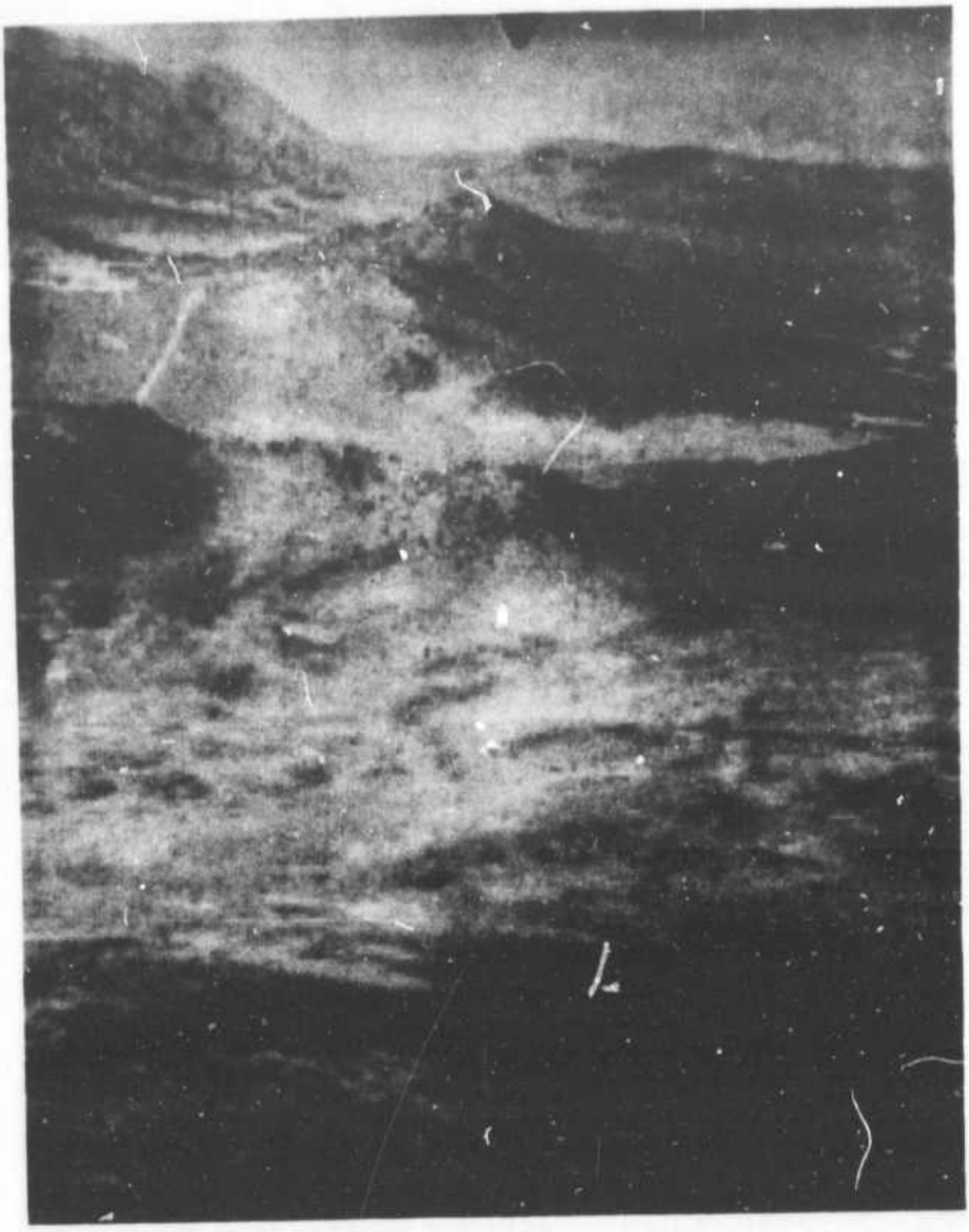


Fig. 13 256x256 Original Image

Fig. 14 256x256 Coded Image, 2 bits/picture element

Fig. 15 256x256 Coded Image, 1 bit/picture element

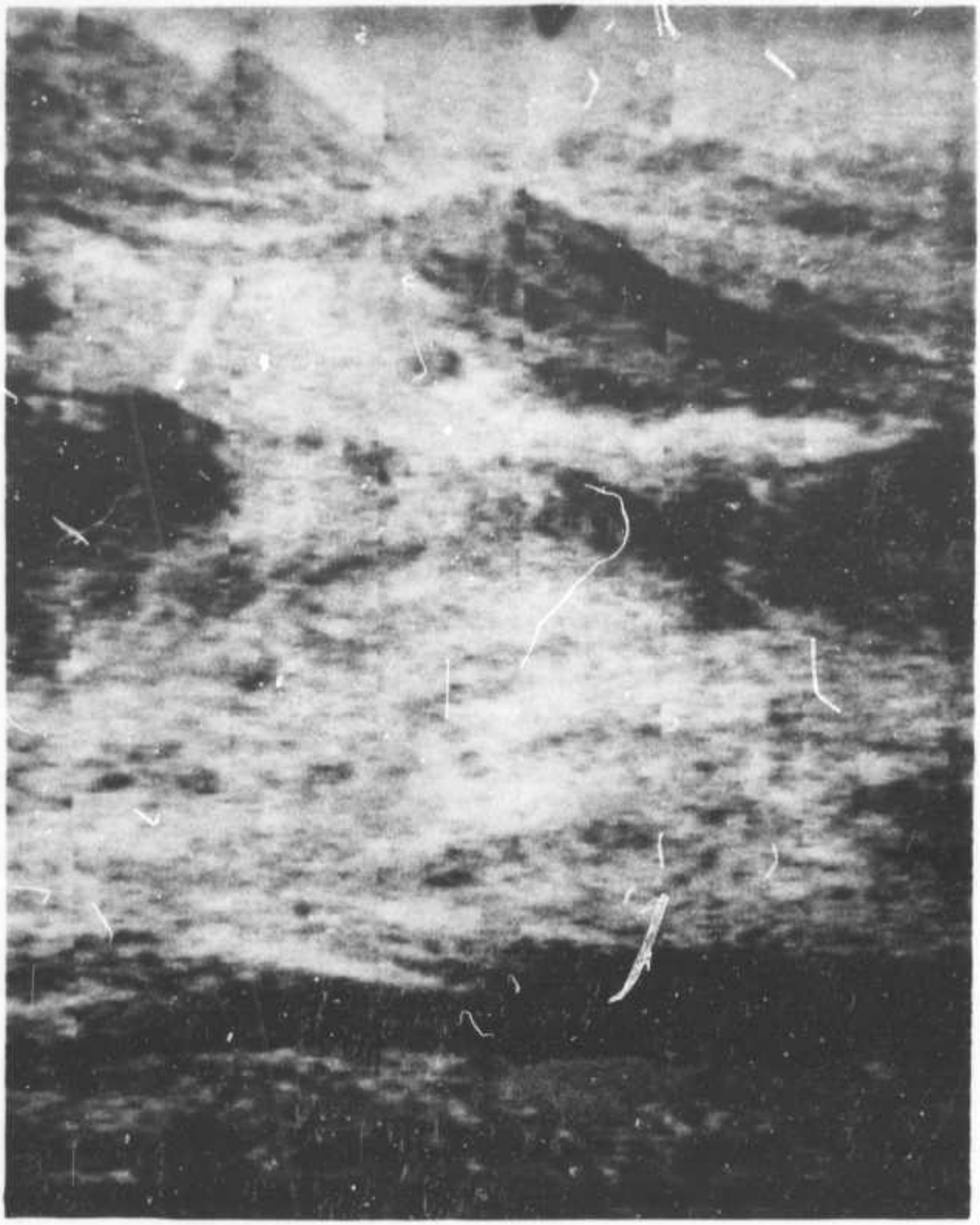


Fig. 16 256x256 Coded Image, 1/2 bit/picture element

EDGE EXTRACTION TECHNIQUES

G. Tang and T. S. Huang

I. INTRODUCTION

The ability of a local operator to extract adequate edge information is important in many applications of image processing. It has been used in image enhancement [1] and image segmentation.

Various local edge operators have been proposed [3], [4], [6], [7] since the advent of computers. Most of them are designed originally for extracting edges from noise-free pictures or nearly noise-free pictures. Based on the philosophy behind these operators, we may classify them into 2 categories; i.e., template-matching operators or difference operator. Template matching operators report the closeness between an ideal edge paradigm and an actual local region on the picture being processed. Difference operators report the difference in brightness.

A general scheme to extract edges from a noisy degraded picture is shown in Fig. 1. The pre-processor is used to reduce the influence of random disturbances to a less significant degree so that the difference operator can report meaningful edge information. Difference operator reports the gray level difference in a certain way. The testor is used to decide whether the iterative application of the difference operator is required and the number of iterations. [7] The post-processor has two functions. The first is to emphasize the edge information reported from the difference operator so that subtle edges will not be missed and to reduce the thickness of reported edges. The second is to change a multiple-level picture to a binary picture by setting a threshold.

In this report we proposed an edge-emphasizing post-processor which is proven to improve the quality of the edges a great deal. We are going to call

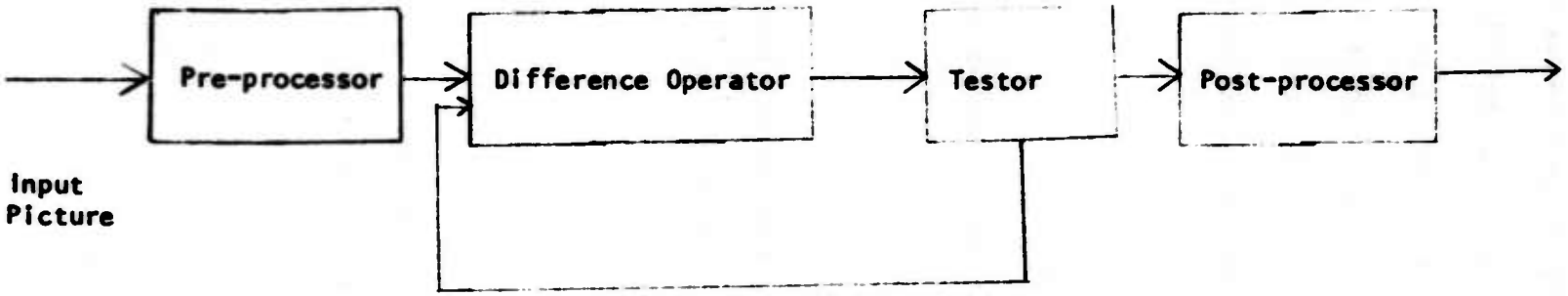


Fig. 1 A general scheme to extract edges from a noisy degraded picture.

this processor "Local unit transformer" which means the transform of gray levels from a global unit to a local unit. We will explain it in the second section. In addition, we test the performance of various difference operators with different post-processors on testing pictures.

Fig. 2 is an original picture without any disturbance or degradation. Fig. 3 is derived from Fig. 2 by adding an artificial random noise with SNR = 6dB. Fig. 4 is obtained from Fig. 2 by replacing the gray level of each point with the average gray level of 25 neighboring points, i.e.,

$$g'(i,j) = \frac{1}{25} \sum_{i-2}^{i+2} \sum_{j-2}^{j+2} g(i,j)$$

Fig. 5 is a noisy degraded picture by adding random noise with SNR = 10dB to Fig. 4.

II. LOCAL UNIT TRANSFORMER

A Local Unit Transformer transforms a picture matrix $g(i,j)$ into another picture matrix $g'(i,j)$ by the following formulae

$$g'(i,j) = g(i,j)/s(i,j)$$

$$s(i,j) = K\left[\sum_{\ell=j-M}^{\ell=j+M} \sum_{K=i-N}^{K=i+N} g^2(K,\ell)\right]^{1/2}$$

N,M represent a small neighborhood around i,j for example 3x3 or 5x5. S(i,j) actually represents the energy of that area defined by N,M. K is a scaling constant.

In order to visualize the idea of a local unit transformer, we can examine it in a simple 1 dimensional case. Assuming we have two 5 dimensional row vectors V1 and V2 as:

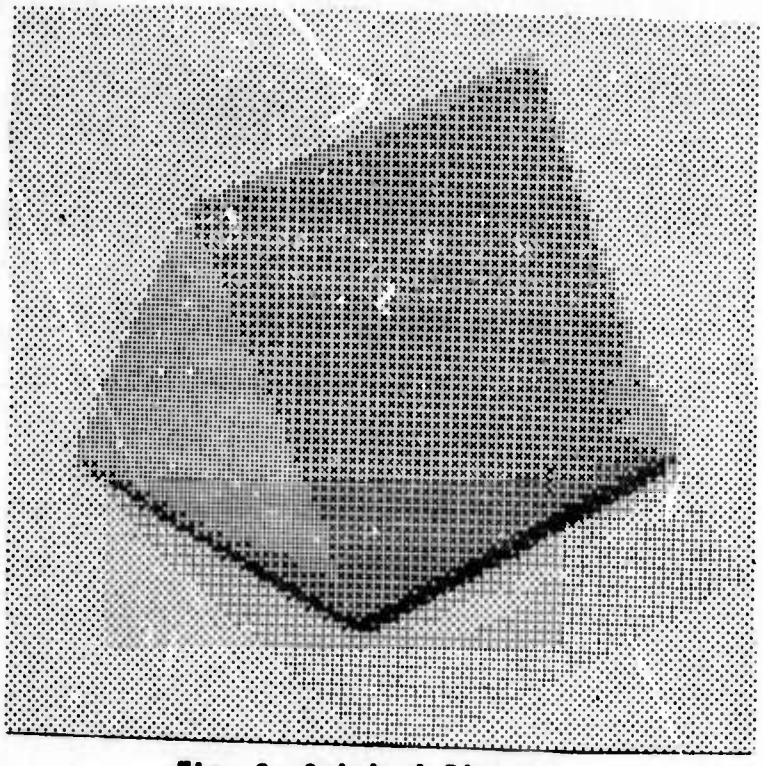


Fig. 2 Original Picture

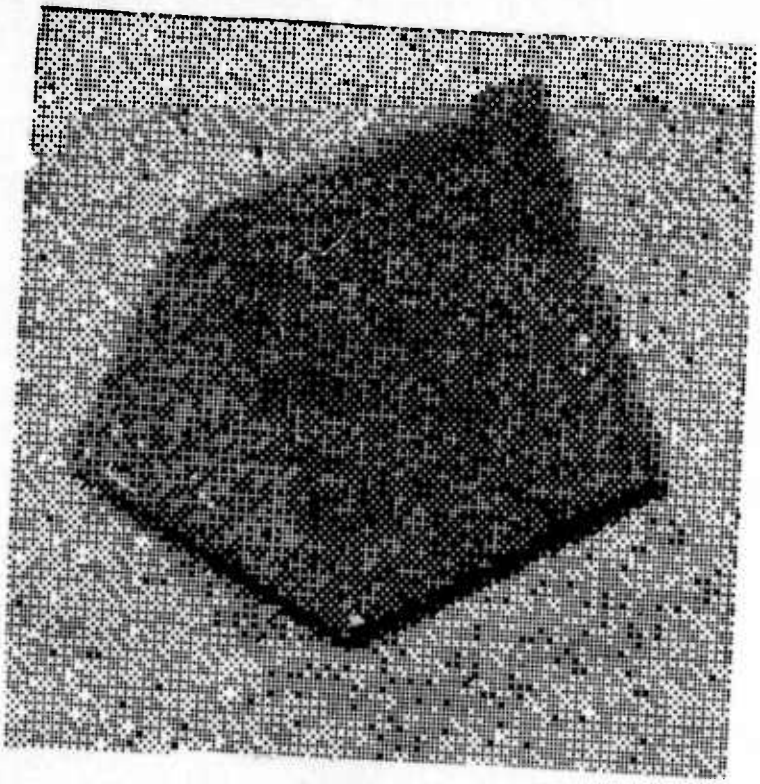


Fig. 3 A noisy picture from Fig. 2 with SNR = 6 dB

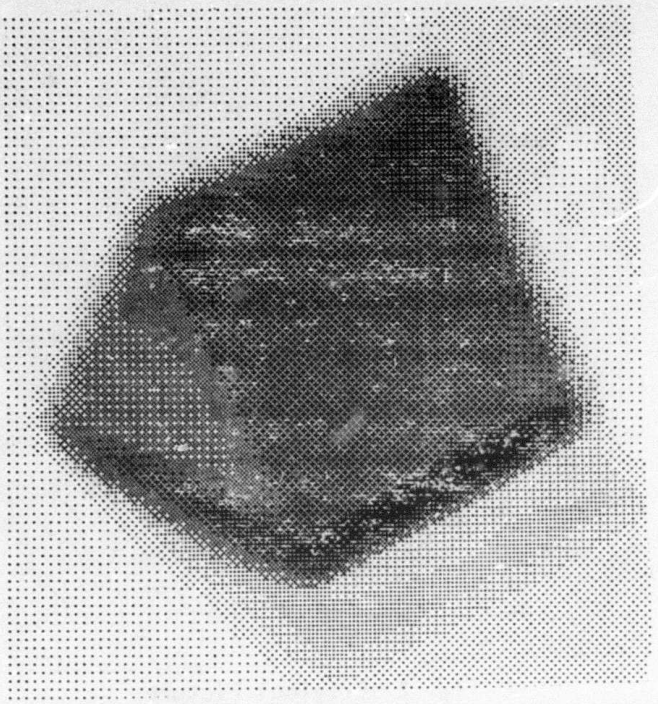


Fig. 4 A degraded picture from Fig. 2 by replacing the gray value at each point with the average of 25 adjacent points. i.e.,

$$g'(i,j) = \frac{1}{25} \sum_{i-2}^{i+2} \sum_{j-2}^{j+2} g(i,j)$$

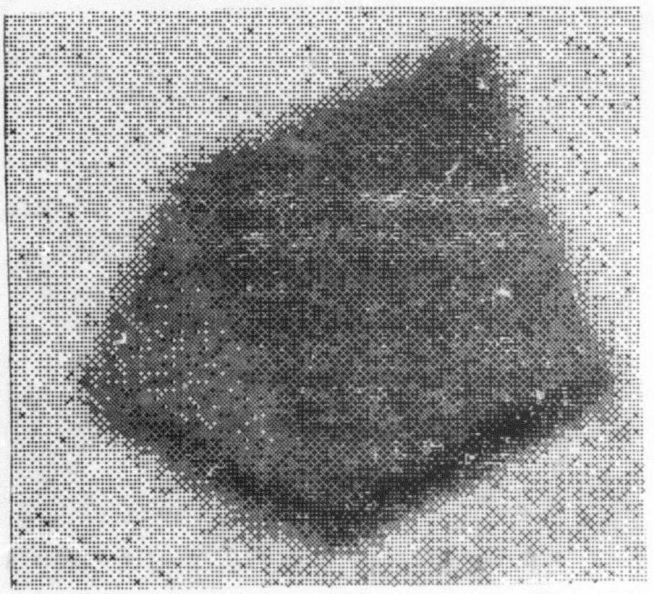


Fig. 5 A noisy picture from Fig. 4 with SNR = 10 dB

$V1 = [1 \quad 1 \quad 4 \quad 10 \quad 10]$

$V2 = [1 \quad 1 \quad 10 \quad 50 \quad 50]$

the outputs of the transformer are $V1'$ and $V2'$ respectively. If we pick K as 1000, we will obtain

$V1' = [67.6 \quad 67.6 \quad 270 \quad 676 \quad 676]$

$V2' = [12.9 \quad 12.9 \quad 129 \quad 645 \quad 645]$

Therefore, the difference between the maximum and minimum components of the output vectors is of about the same order.

For further illustration, we may consider another two vectors:

$V1 = [1 \quad 1 \quad 10 \quad 1 \quad 1]$

$V2 = [1 \quad 1 \quad 100 \quad 1 \quad 1]$ with K=1

so,

$$V1' = \left[\frac{1}{\sqrt{104}}, \frac{1}{\sqrt{104}}, \frac{10}{\sqrt{104}}, \frac{10}{\sqrt{104}}, \frac{1}{\sqrt{104}}\right]$$

$$V2' = \left[\frac{1}{\sqrt{10004}} \quad \frac{1}{\sqrt{10004}} \quad \frac{100}{\sqrt{10004}} \quad \frac{1}{\sqrt{10004}} \quad \frac{1}{\sqrt{10004}}\right]$$

$V1' \doteq [0 \quad 0 \quad 1 \quad 0 \quad 0]$

$V2' \doteq [0 \quad 0 \quad 1 \quad 0 \quad 0]$

Again, the difference between the maximum and the minimum is of the same order.

The examples also indicate that the order of the applications of the transformer and the difference operator ought not to give any significantly different results as far as edge detection is concerned.

In practice we might set up a lower bound (LA) for the local energy s(I,J). The transformer is then modified as

$$g'(I,J) = g(I,J)/s(I,J) \quad \text{if } s(I,J) \geq LA$$

$$= 0 \quad \text{if } s(I,J) < LA$$

where s(I,J) is defined as

$$s(I,J) = K\left[\sum_{\ell=J-M}^{J+M} \sum_{K=I-M}^{I+M} g^2(K,\ell)\right]^{1/2}$$

III. LOCAL OPERATORS AND THEIR DESIGNATIONS

In this section we define several local operators and associate each with a proper designation.

(I) Local Unit Transformer

Definition as in the previous section.

Designation: T[2N+1, 2M+1, K, LA]

(II) Robert's Cross Operator

Definition: $g'(I,J) = |g(I-N, J-M) - g(I+N, J+M)|$

$$+ |g'(I-N, J+M) - g(I+N, J-M)|$$

Designation: R[2N+1, 2M+1]

(III) Lapiacian Operator

Definition: $g'(I,J) = |g(I-N, J-M) + g(I+N, J+M)$

$$+ g(I-N, J+M) + g(I+N, J-M) - 4g(I,J)|$$

Designation: L[2N+1, 2M+1]

(iv) E Operator

Definition: $g'(i,j) = [g(i,j) - \overline{g(i,j)}]^+$

$[\cdot]^+$ represents the positive value of $[\cdot]$, i.e., if $(\cdot) \geq 0$, then $[\cdot]^+ = (\cdot)$ if $(\cdot) < 0$, then $[\cdot]^+ = 0$

and $$\overline{g(i,j)} = \frac{1}{(2N+1)(2M+1)} \sum_{\ell=i-N}^{\ell=i+N} \sum_{k=j-M}^{k=j+M} g(\ell,k)$$

Designation: E[2N+1, 2M+1]

(v) S Operator

Definition: $g'(i,j) = \max (S_1(i,j) , S_2(i,j)$

$$S_1(i,j) = 2Sa - Sb - Sc$$

$$S_2(i,j) = 2Sa' - Sb' - Sc'$$

where

$$Sa = \sum_{\ell=j-M}^{j+M} g(i,\ell)$$

$$Sb = \sum_{\ell=j-M}^{j+M} g(i-N,\ell)$$

$$Sc = \sum_{\ell=j-M}^{j+M} g(i+N,\ell)$$

$$Sa' = \sum_{\ell=j-N}^{j+N} g(\ell,j)$$

$$Sb' = \sum_{\ell=j-N}^{j+N} g(\ell,j-M)$$

$$Sc' = \sum_{\ell=j-N}^{j+N} g(\ell,j+M)$$

Designation: S[2N+1, 2M+1]

(vi) G Operator

Definition: $$g'(i,j) = \left| \sum_{\ell=i-N}^{i-1} g(\ell,j) - \sum_{\ell=i+1}^{i+N} g(\ell,j) \right| + \left| \sum_{k=j-M}^{j-1} g(i,k) - \sum_{k=j+1}^{j+M} g(i,k) \right|$$

Designation: G[2N+1, 2M+1]

(vii) B Operator

Definition: Let D(k) equal the number of picture elements with gray level k. P is a present percentage. ITHRESH is the smallest gray level such that $\sum_{ITHRESH}^{IMAX} D(k) \geq p \times X$ (X is the total number of picture element), IMAX is the maximum gray level.

Then, we have

$$g'(i,j) = 1 \text{ if } g(i,j) \geq ITHRESH$$

$$= 0 \text{ otherwise.}$$

Designation: B[P]

In addition to the notations for operators, we use P1 to stand for Fig. 2; P2, Fig. 3; P3, Fig. 4; P4, Fig. 5. Thus, we use the notation B[5] S(3,3) E(3,3) P1 to stand for the picture obtained by the following sequence of operations on Fig. 2. First we apply E(3,3) on Fig. 2 and get a new picture P1', then we apply S(3,3) on P1' and get P1''. Finally we apply B(5) to P1'' and get P1'''. B[5] S[3,3] E(3,3) P1 represents P1'''.

IV. EXPERIMENT

A picture of a pyramid is digitized into a square matrix of size 128 by 128. The gray level at each picture element ranges from 0 to 255. These pictures are stored on magnetic tapes. Since all the operators except B operator compress the input picture, we add some artificial rows and columns to the

picture matrix so as to retain the size of output picture to be 128x128. For example, in the case of G(5,5) P1, the top 2 rows of the output matrix is made artificially to be the same as the third row; the lowest two rows equal the lowest third row; the first two columns equal the third column; the last two columns equal the last third column.

Because the limitations on the electrostatic printer, the output picture is rescaled with 17 gray levels and displayed in the fixed-pattern mode. The experimental results are shown in Figs. 6-35.

V. CONCLUSION

Experimental results indicate that the application of the transformer indeed improves the quality of the output edges quite a bit. For instance, Fig. 13 and Fig. 14 have almost the same number of picture elements. But Fig. 14 contains much more meaningful edges than Fig. 13 does. Even some edges between the shaded area and background(where the differences in gray level is fairly small in comparison with others)can be seen on Fig. 14.

Also it is found that the gradient operator with an appropriate post-processor can give a set of edges which retain most of the edge information on the original picture.

In this report, we are dealing mainly with the behavior of local operators. Other questions such as continuity of edge points, separation between relevant and irrelevant edges, and descriptions of edges are not considered.

REFERENCES

[1] T. S. Huang and G. Y. Tang, "Application of Edge Detection to Image Enhancement," TR-EE 75-2, School of Electrical Engineering, Purdue University, West Lafayette, Indiana, November 1974.

[2] M. Y. Yoo and T. S. Huang, "Noise Reduction in Photographic Images, " this report.

[3] A. K. Griffith, "Mathematical Models for Automatic Line Detection," Commun. Assn. Comput. Mach., Vol. 20, No. 1, pp. 62-80, January 1973.

[4] M. H. Hueckel, "An Operator Which Locates Edges in Digitized Picture," Journal of the Assn for Comp. Machine, Vol. 18, No. 1, pp. 113-125, January 1971.

[5] R. O. Duda and P. E. Hart, Pattern Classification and Scene Analysis, Wiley-Interscience Publication, pp. 263-297, 328-335.

[6] L. G. Roberts, "Machine Perception of Three-Dimensional Solids," in Optical and Electro-optical Information Processing, J. T. Tippett et al, eds., (MIT Press, Cambridge, Mass.,) pp. 159-197, 1965.

[7] T. Kasvand, "Iterative Edge Detection," to appear in Computer Graphics and Image Processing.

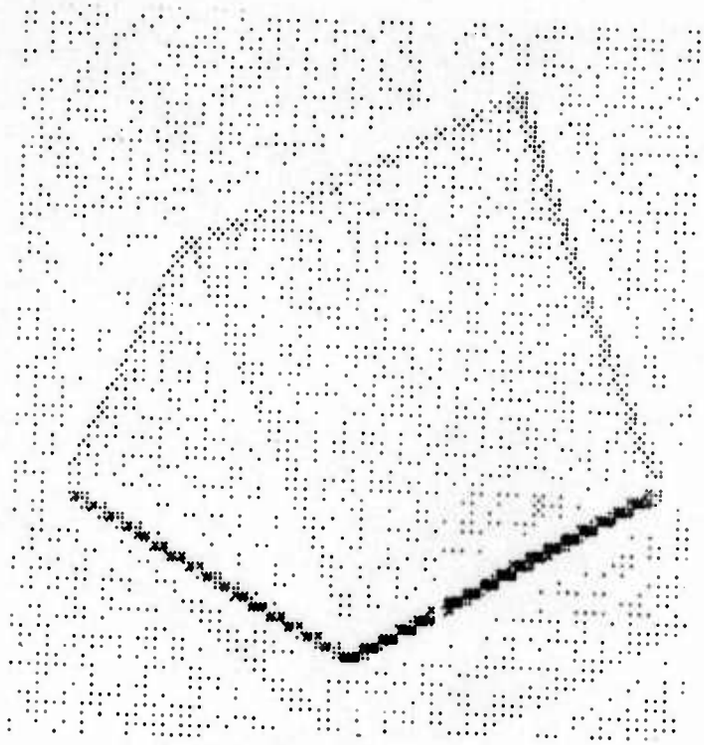


Fig. 6 E[3,3] P1

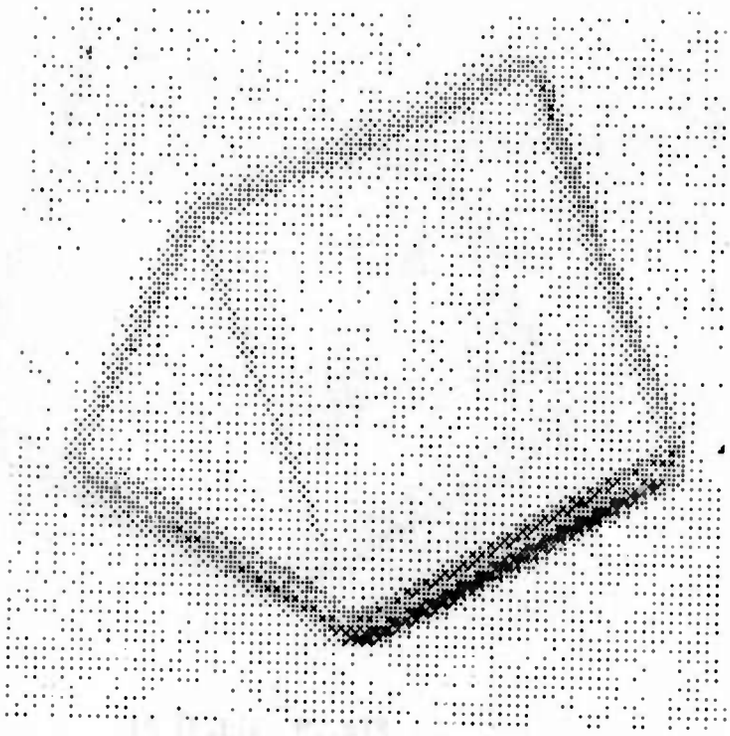


Fig. 7 R[3,3] P1

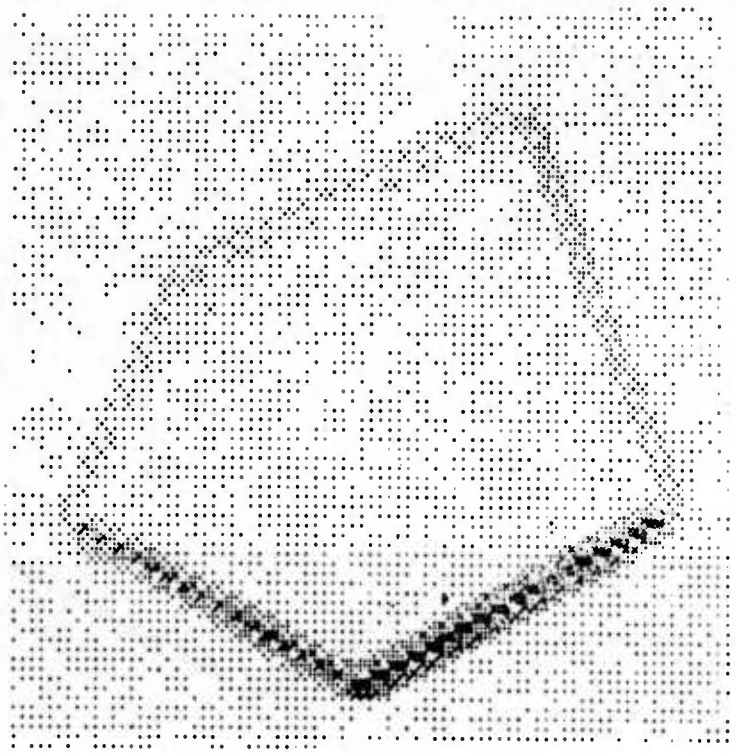


Fig. 8 L[3,3] P1

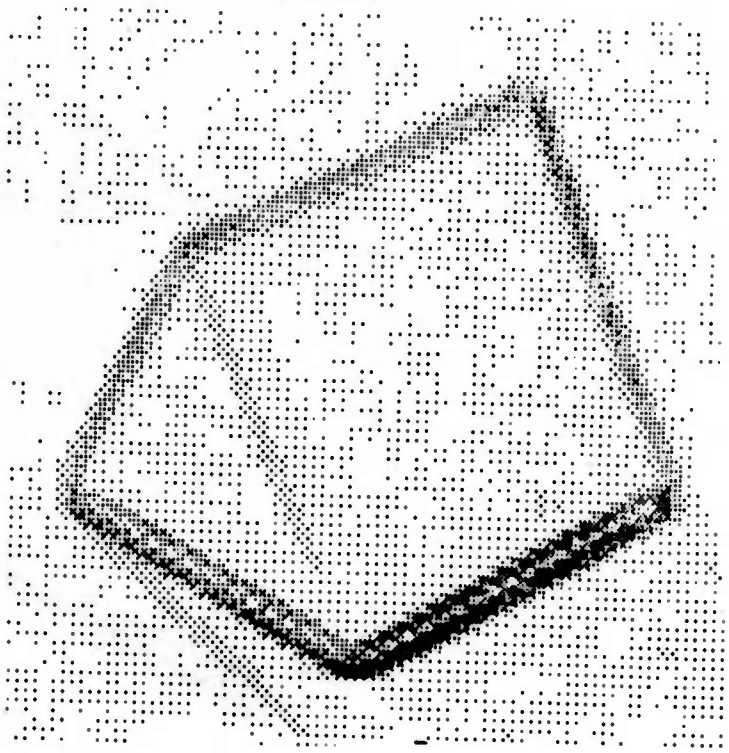


Fig. 9 G[3,3] P1

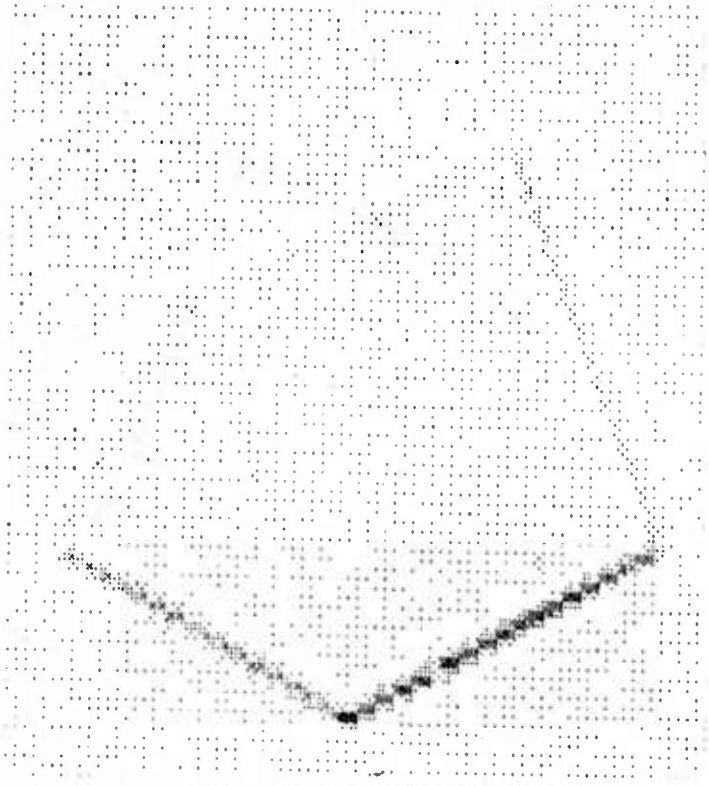


Fig. 10 S[3,3] E[3,3] P1

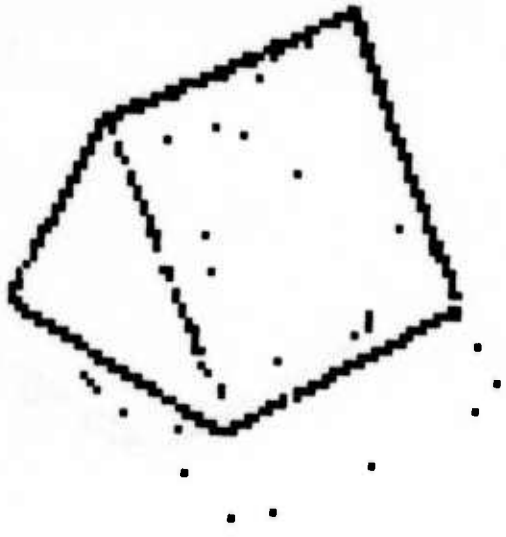


Fig. 11 B[3.5]E[3,3] P1

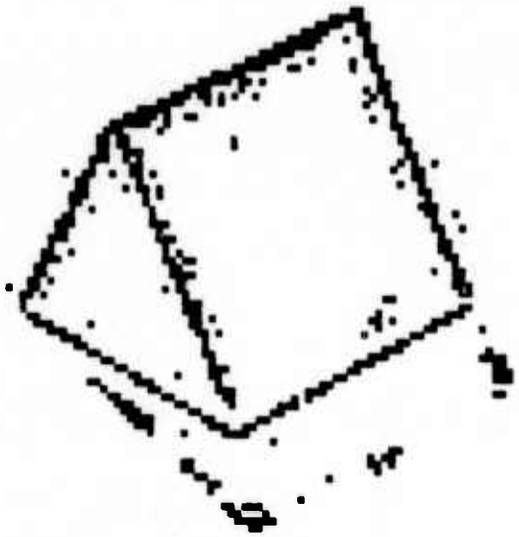


Fig. 12 B[4] T[5,5,1000,5] E[3,3] P1

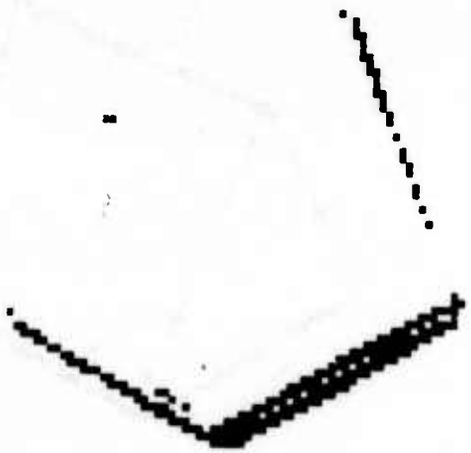


Fig. 13 B[3,5] R[3,3] P1

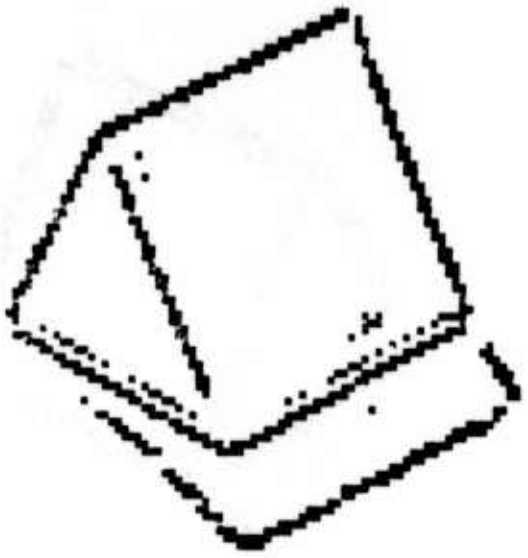


Fig. 14 B[4] T[5,5,1000,5] R[3,3] P1

Fig. 15 B[3.5] L[3,3] P1

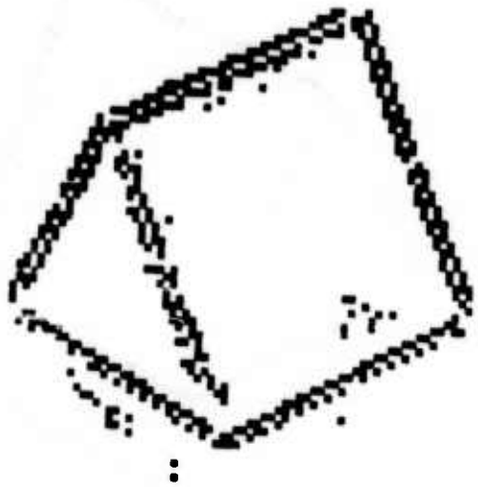


Fig. 16 B[4] T[5,5,1000,5] L[3,3] P1

Fig. 17 B[3.5] G[3,3] P1

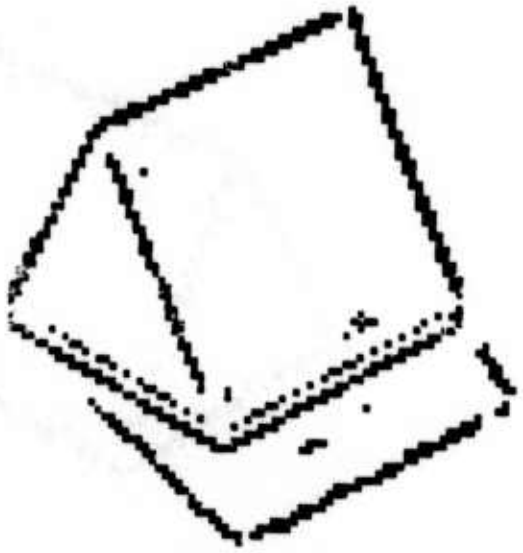


Fig. 18 B[4] T[5,5,1000,5] L[3,3] P1

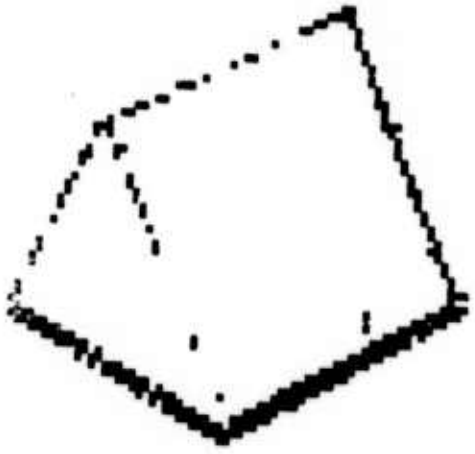


Fig. 19 B[3.5] S[3,3] E[3,3] P1

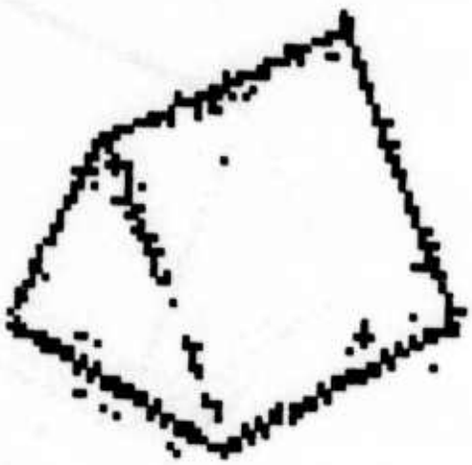


Fig. 20 B[4] T[5,5,1000,5] L[3,3] E[3,3] P1

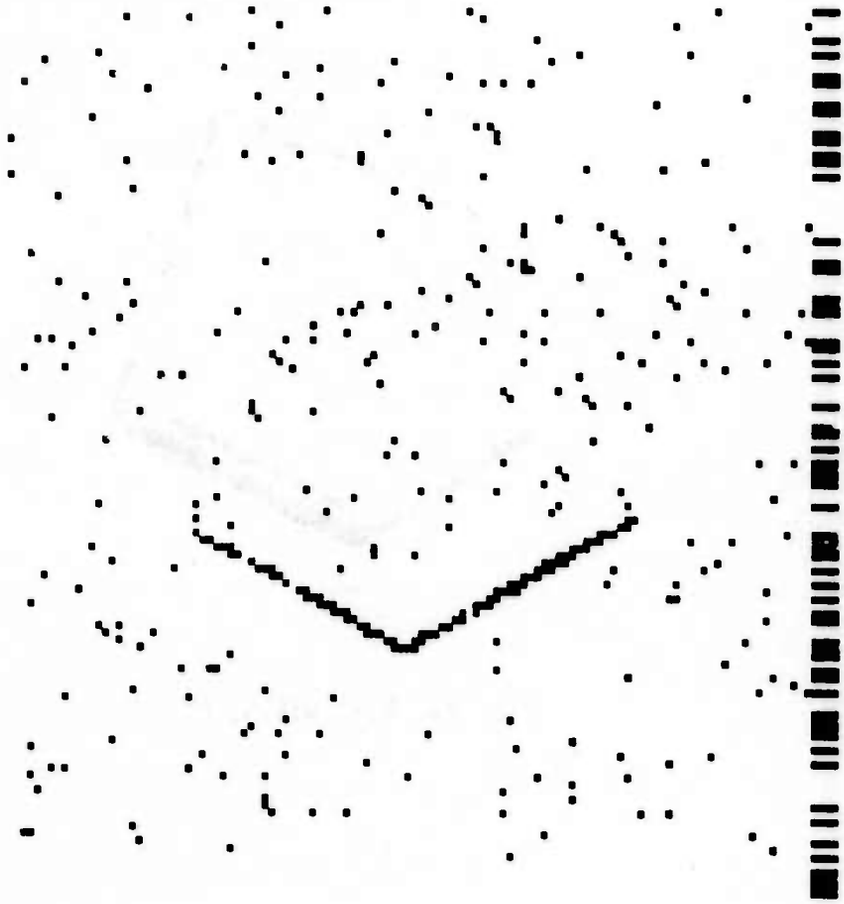


Fig. 21 B[4] E[5,5] P2

Fig. 22 B[4] G[5,5] P2

Fig. 23 B[4] T[5,5,1000,150] G[5,5] P2

Fig. 24 B[6] T[5,5,1000,150] G[5,5] P2

Fig. 25 B[10] T[5,5,1000,150] G[5,5] P2

Fig. 26 B[5] S[5,5] E[5,5] P3

Fig. 27 B[5] R[5,5] P3

Fig. 28 B[5] L[5,5] P3

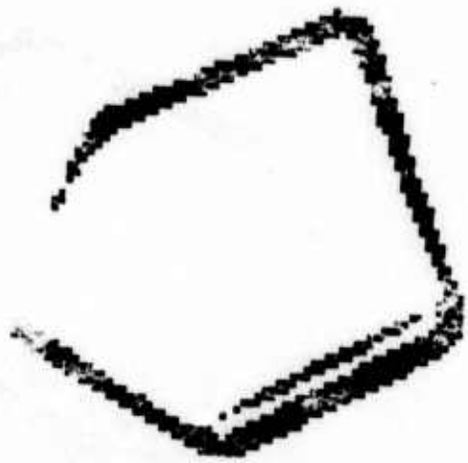


Fig. 29 B[5] G[5,5] P3

Fig. 30 B[5] T[5,5,1000,25] R[5,5] P3

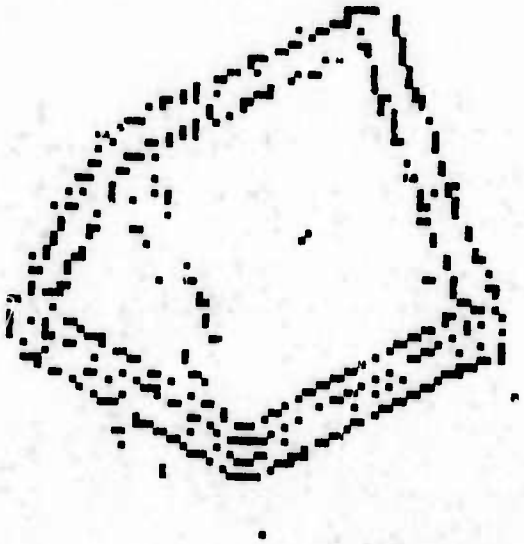


Fig. 31 B]5] T[5,5,1000,25] L[5,5] P3

Fig. 32 B[5] T[5,5,1000,25] G[5,5] P3

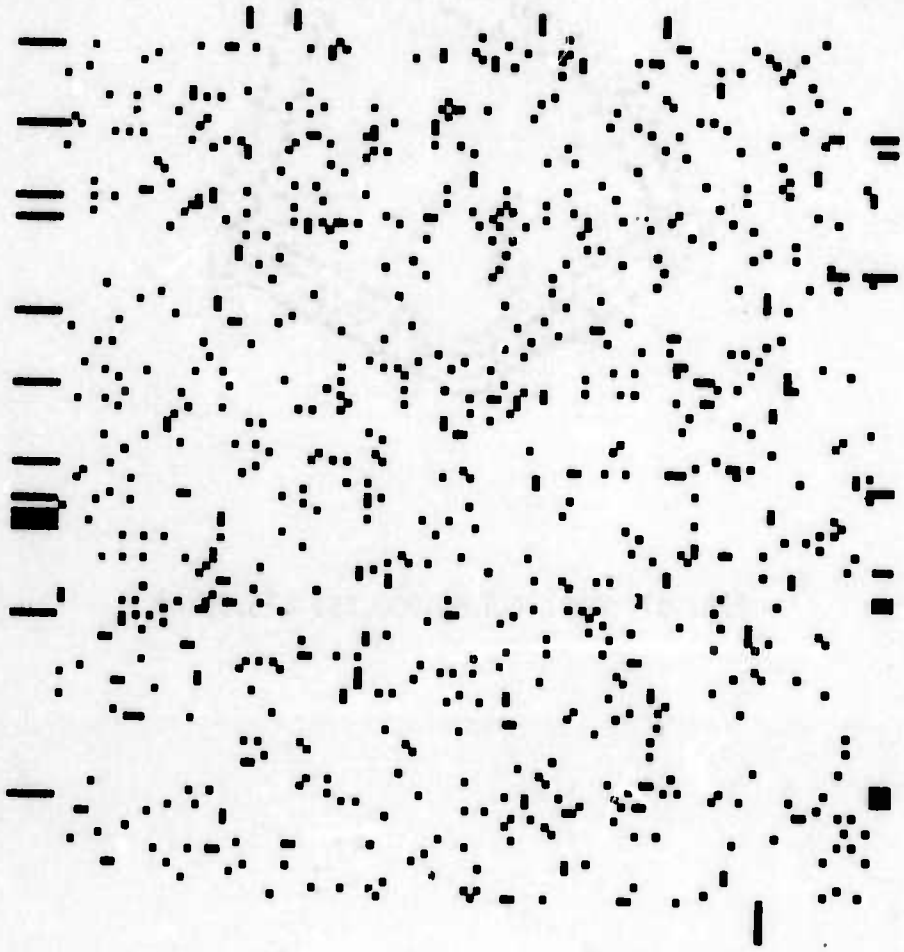


Fig. 33 B[6] L[5,5] E[5,5] P4

Fig. 34 B[6] G[5,5] P4

Fig. 35 B[6] T[5,5,1000,120]G[5,5] P4

IMAGE DECOMPOSITION

J. Burnett and T. S. Huang

A scan line of a simple picture is modeled as the sum of two random processes. A Markov jump process is used to model abrupt changes in grey level due to boundaries. A second random process represents variation about the average gray level due to noise or texture. The optimum mean square filter is implemented to extract the boundary process. This estimate of the boundary process is then differentiated to locate the jumps or boundary locations. A comparison is made with the best linear techniques for performing the same operations.

Many image processing tasks such as restoration, pattern classification and data compression can be performed with better results than otherwise possible if the image is first segmented or partitioned into regions with common properties. Nahi and Habibi [1] and Huang and Tang [2] have suggested this technique for image restoration, Gupta [4] and Robertson [3] for pattern classification and Gupta [4] for data compression

One common approach to image segmentation is to take the gradient or laplacian of the picture to find the boundaries. However, if there is much variation in gray level at non boundary points due to noise or texture in the image this technique does not work well.

To overcome this difficulty we model an image $I(X,Y)$ as being the sum of two processes: A Markov jump process (i.e., a process whose values are piecewise constant) $J(X,Y)$ which models the discontinuities in average gray levels due to boundaries; and a second process $T(X,Y)$ (where $T(X,Y) = I(X,Y) - J(X,Y)$) where $T(X,Y)$ represents texture or noise in the image. A sample function of a one-dimensional jump process with two possible values is shown in Fig. 1.

With this model the problem of boundary finding can be stated as follows:

given $I(X,Y) = J(X,Y) + T(X,Y)$

what is the best estimate $\hat{J}(X,Y)$ of $J(X,Y)$?

Once we have $\hat{J}$ we can presumably differentiate it to obtain the locations of the jumps or boundaries.

In general finding $\hat{J}$ is a very difficult problem. However, some one-dimensional results have been obtained [5], [6]. Let $X(t)$ be a Markov jump process with possible values $a_1, a_2, \ldots a_k$. We assume

$$(1)\quad P_{ij}(h) = P(X(t+h) = a_j \mid X(t) = a_i) = \begin{cases} 1 - v_i h & j = i \\ v_{ij} h & j \neq i \end{cases} \qquad (1)$$

where v_{ij} is a constant and

$$v_i = \sum_{\substack{j=1 \\ j \neq i}}^{k} v_{ij}$$

Let the observations be

$$(2)\quad Y(t) = X(t) + BU(t) \text{ where} \qquad (2)$$

$U(t)$ is white noise.

Let $P_j(t) = P(X(t) = a_j \mid Y(S)\ 0 \leq S \leq t)$

Then P_j obeys the differential equation

$$\frac{dP_j(t)}{dt} = -v_j P_j(t) + \sum_{\substack{i=1 \\ i \neq j}}^{k} v_{ij} P_i(t) - B^{-2}\,\overline{X}\,[a_j - \overline{X}]P_j(t) + B^{-2}[a_j - \overline{X}] \qquad (3)$$

where $\overline{X} = \Sigma\, a_k\, P(X = a_k \mid Y(S)\ 0 \leq S \leq t)$

In particular we consider the special case of a random telegraph wave [7]. This would be a good model for a scan line of a picture with two average gray

levels but some variation about each level due to texture or noise. In the above formulae we have $k = 2$, $a_1 = 1$, $a_2 = -1$, v = expected number of jumps per unit time = 1,

$$\hat{X} = \text{minimum mean square estimate of } X$$
$$= E(X(t) \mid Y(S)\ 0 \leq S \leq t) = 1 \cdot P_1(t) - 1 \cdot P_2(t) \tag{4}$$

From (3) and (4) we have

$$\dot{\hat{X}} = -2v\,\hat{X} - B^{-2}X(1-\hat{X}^2) + B^{-2}(1-\hat{X}^2)\,Y(t) \tag{5}$$

A sample of a random telegraph wave was generated and is shown in Fig. 1. Noise was added and the result is shown in Fig. 2. The optimum mean square filter (5) was implemented and the output shown in Fig. 3. The output of the filter was differentiated to locate the boundaries and the results are in Fig. 4. For comparison the noisy wave was passed through a Wiener filter (which represents the best <u>linear</u> least square estimate) and the output shown in Fig. 5. Note how much smoothing of the signal has taken place. This corresponds to the smearing mentioned in [1] and is typical of linear estimators. The output of the linear filter was differentiated to yield Fig. 6. The boundary spikes of the optimum filter are both higher and narrower than the linear filter indicating that they would give more accurate estimates of boundary location.

In conclusion we have shown that the modeling of images with jump processes offers the possibility of developing nonlinear processing techniques that will give better results than those obtained by linear methods.

References

[1] N. Nahi and A. Habibi, "Decision Directed Image Enhancement," IEEE Trans. on Circuits and Systems, Vol. CAS 22, March 1975.

[2] T. Huang and G. Tang, "Application of Edge Detection to Image Enhancement in Image Analysis and Modeling," TR-EE 75-2, School of Electrical Engineering, Purdue University, November 1974.

[3] Robertson, Fu and Swain, "Multispectral Image Partitioning, LARS Information Note #071373, Purdue University, 1973.

[4] Jai Gupta, "Multi-Image Modeling, TR-EE 74-24, School of Electrical Engineering, Purdue University, September 1974.

[5] W. Wonham, "Some Applications of Stochastic Differential Equations to Nonlinear Filtering, J. Siam Control, Vol. 2, No. 3, 1964.

[6] P. Frost and T. Kailath, "An Innovations Approach to Least Squares Estimation-Part III: Nonlinear Estimation in White Gaussian Noise," IEEE Trans. Auto. Control, Vol. AC-16, June 1971.

[7] Papoulis, Probability, Random Variables and Stochastic Processes, McGraw Hill, New York 1965.

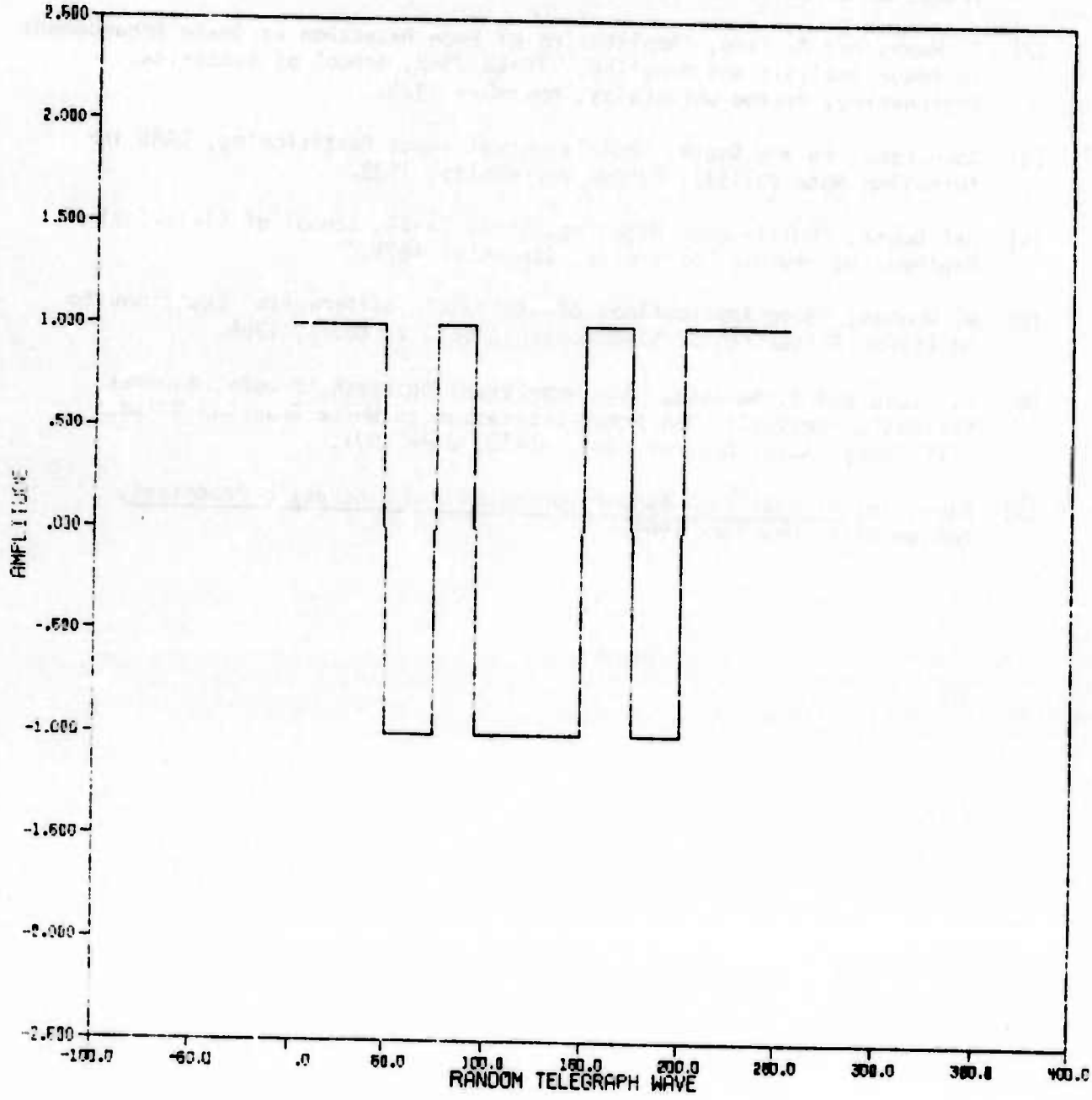


Fig. 1 Random Telegraph Wave

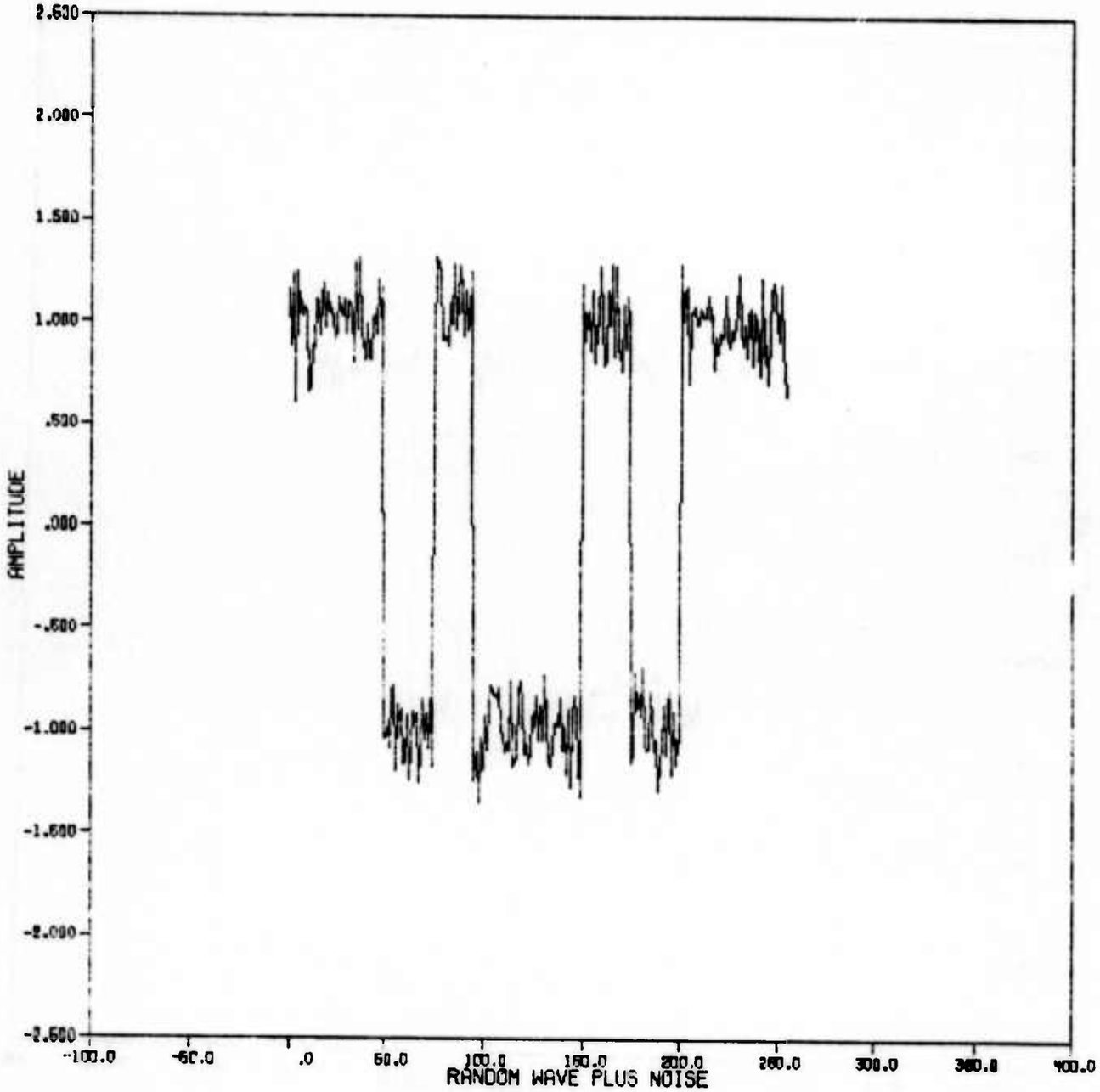


Fig. 2 Random Wave Plus Noise

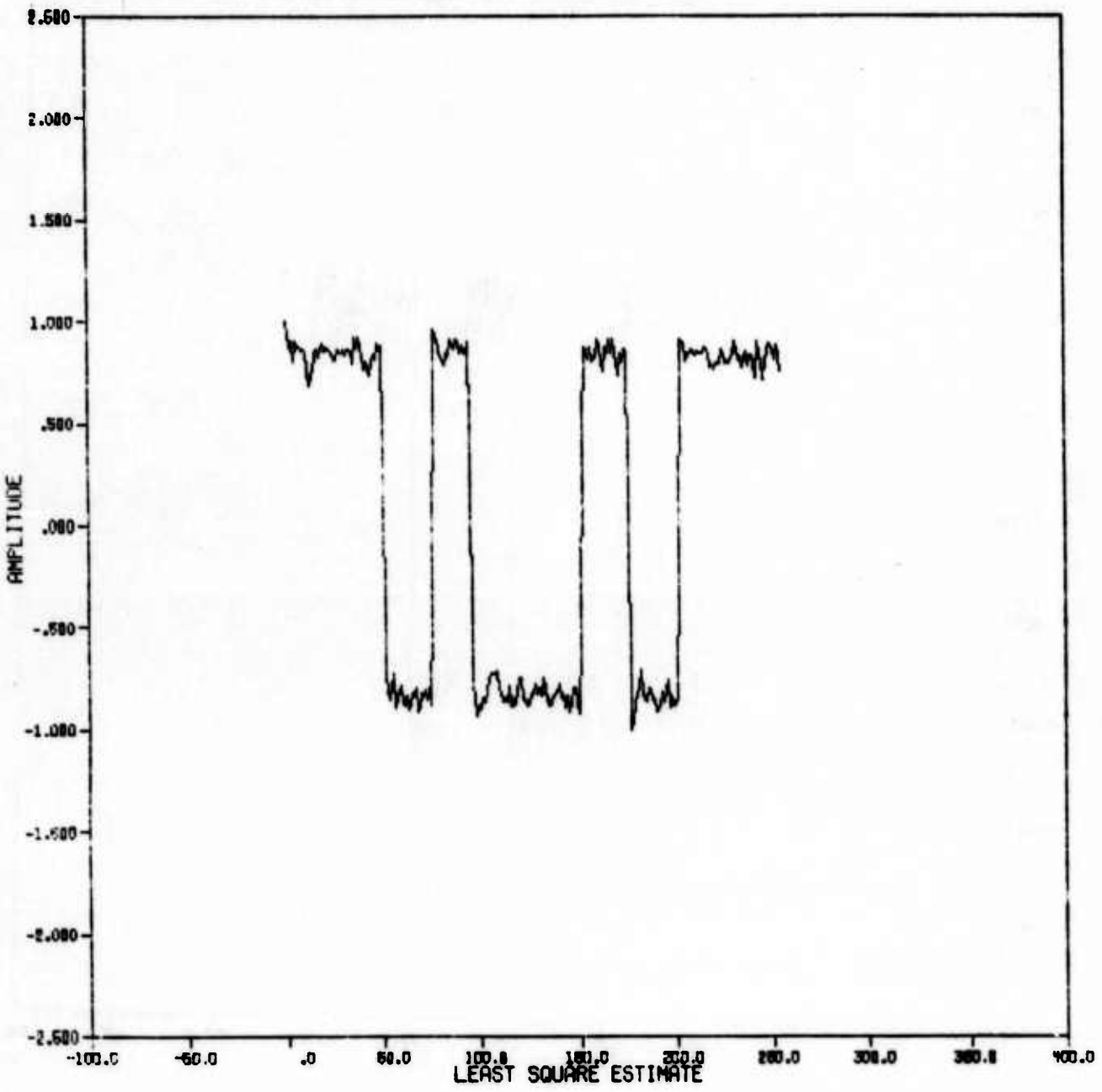


Fig. 3 Least Square Estimate of the Random Telegraph Wave

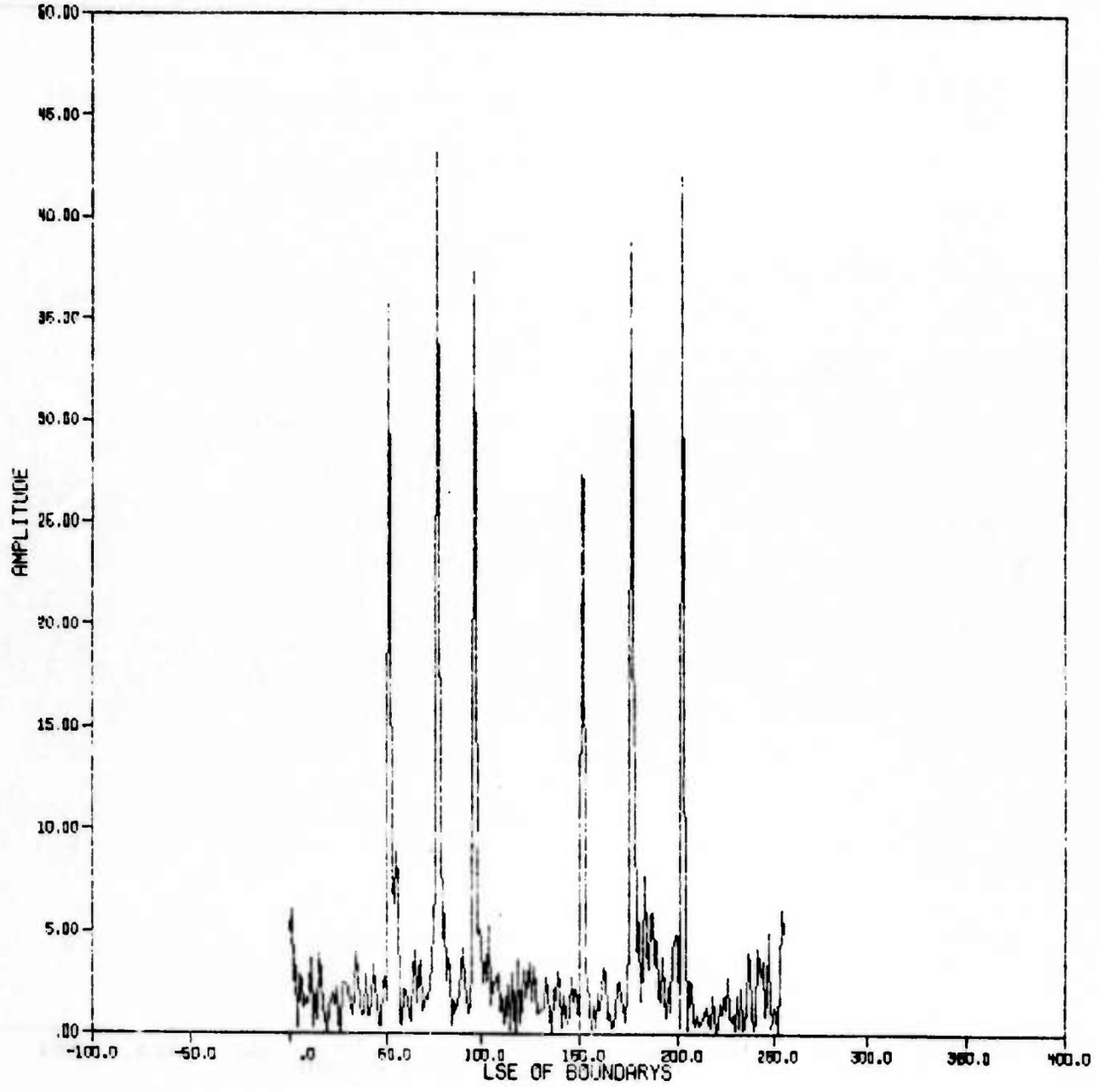


Fig. 4 Least Square Estimate of Boundary Locations

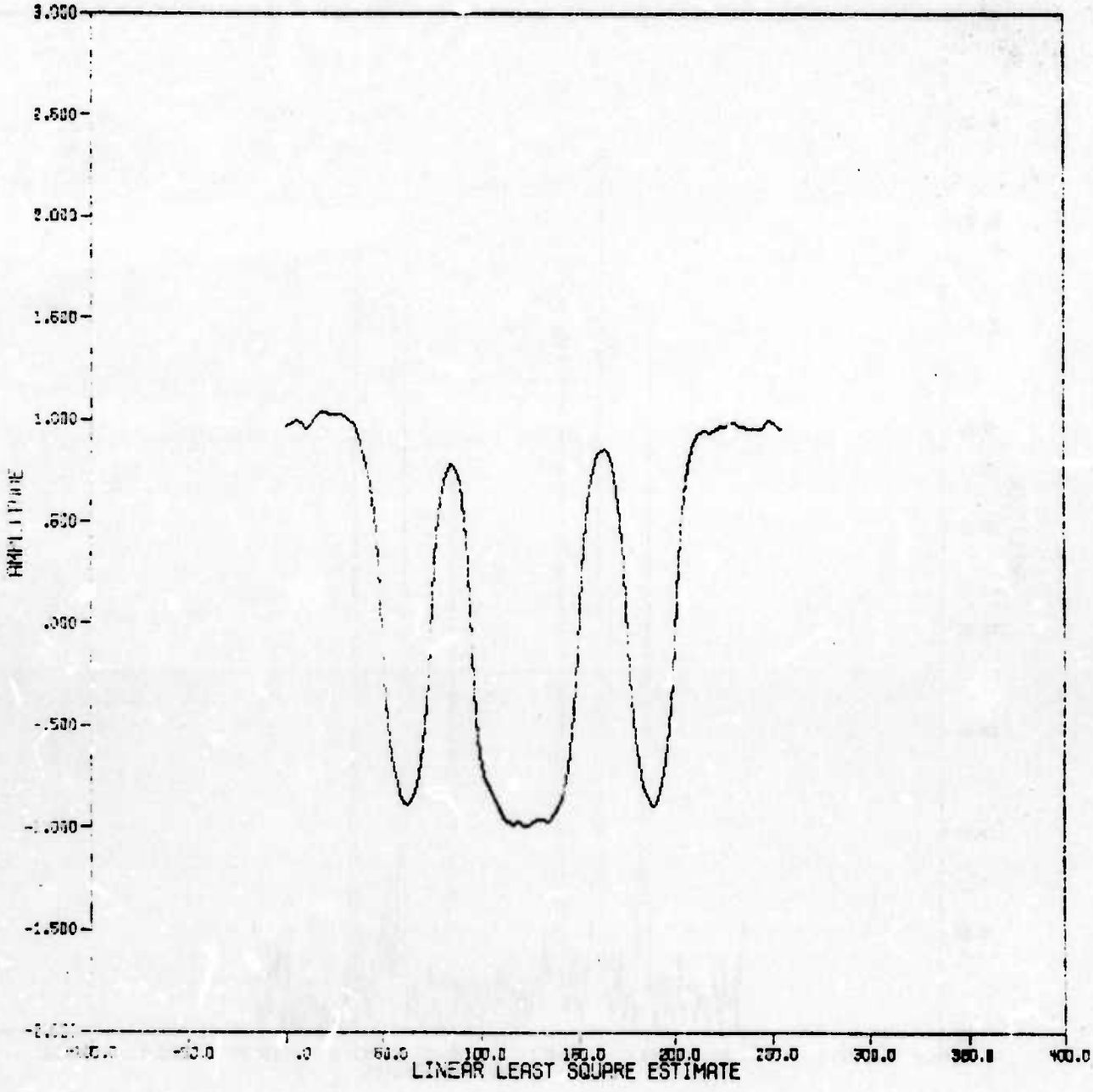


Fig. 5 Linear Least Square Estimate of the Rnadom Telegraph Wave

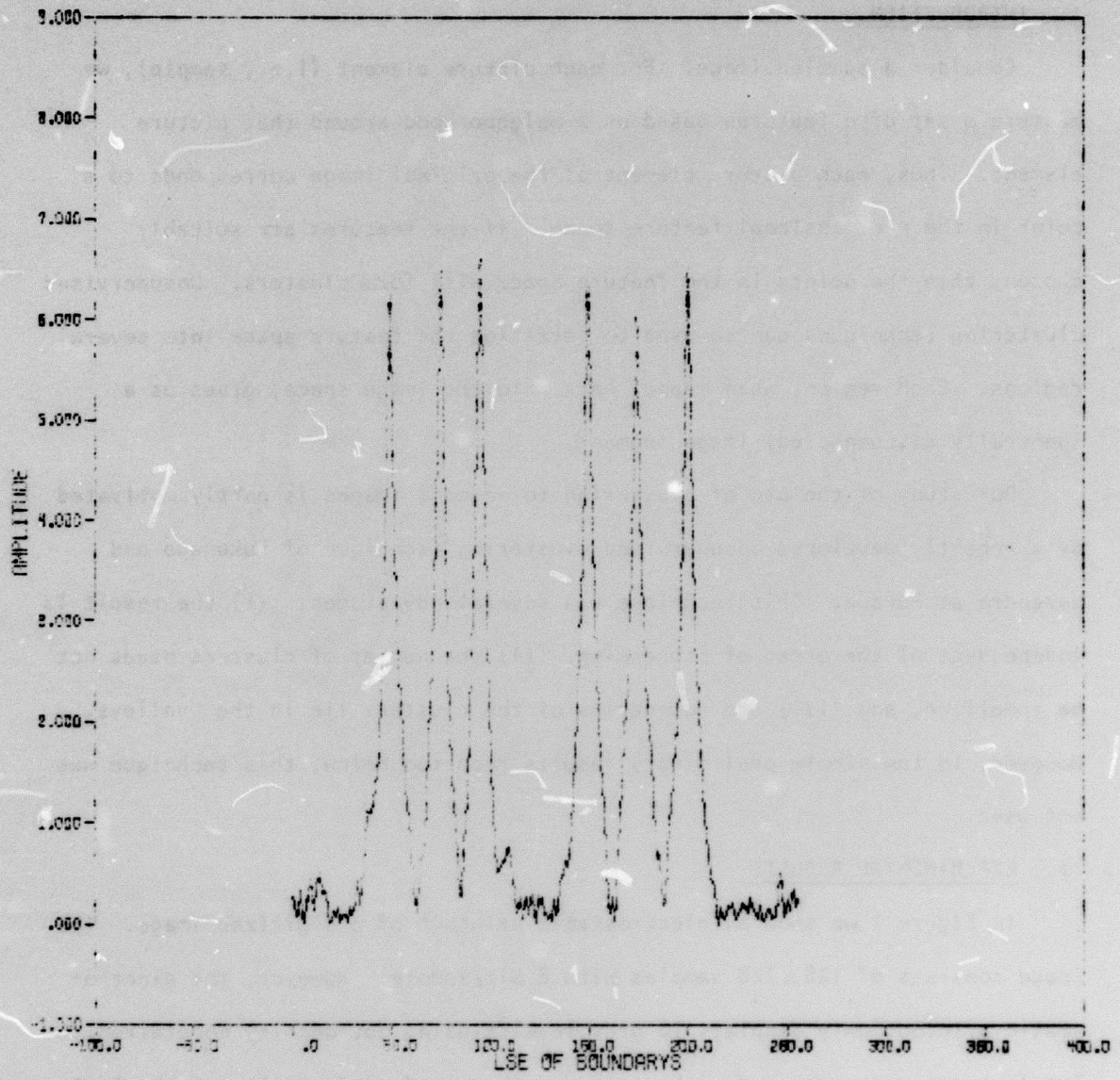


Fig. 6 Linear Least Square Estimate of the Boundary Locations

IMAGE SEGMENTATION BY UNSUPERVISED CLUSTERING

M. Y. Yoo and T. S. Huang

I. INTRODUCTION

Consider a sampled image. For each picture element (i.e., sample), we measure a set of n features based on a neighborhood around that picture element. Thus, each picture element of the original image corresponds to a point in the n-dimensional feature space. If the features are suitably chosen, then the points in the feature space will form clusters. Unsupervised clustering techniques can be used to partition the feature space into several regions. Each region, when mapped back into the image space, gives us a (generally disconnected) image segment.

Our study of the use of clustering to segment images is partly motivated by a recently developed unsupervised clustering technique of Fukunaga and Narendra at Purdue. This technique has several advantages: (i) the result is independent of the order of processing, (ii) the number of clusters needs not be specified, and (iii) the boundaries of the clusters lie in the "valleys." However, in the simple preliminary results reported below, this technique was not used.

II. EXPERIMENTAL RESULTS

In Figure 1 we show an electrostatic printout of a digitized image. The image consists of 128 x 128 samples with 8 bit/sample. However, the electrostatic printout only displays 16 gray levels (using dot-density modulation). Two features were measured: the mean and the standard deviation of the 3 x 3 neighborhood of each picture element. Figure 2 shows the points in the feature space. In Figures 4-5, some results of clustering by eyeballing are shown. In part (b) of each of these Figures, we show in black the image

segment corresponding to the boxed region in the feature space in part (a). The symbols used in parts (a) of Figures 4-6 are explained in Figure 3.

III. FUTURE WORK

We plan to work on the segmentation of multispectral images using the Fukunaga-Narendra valley-seeking unsupervised clustering technique. We would also investigate ways of incorporating spatial information (in image space) into this approach.

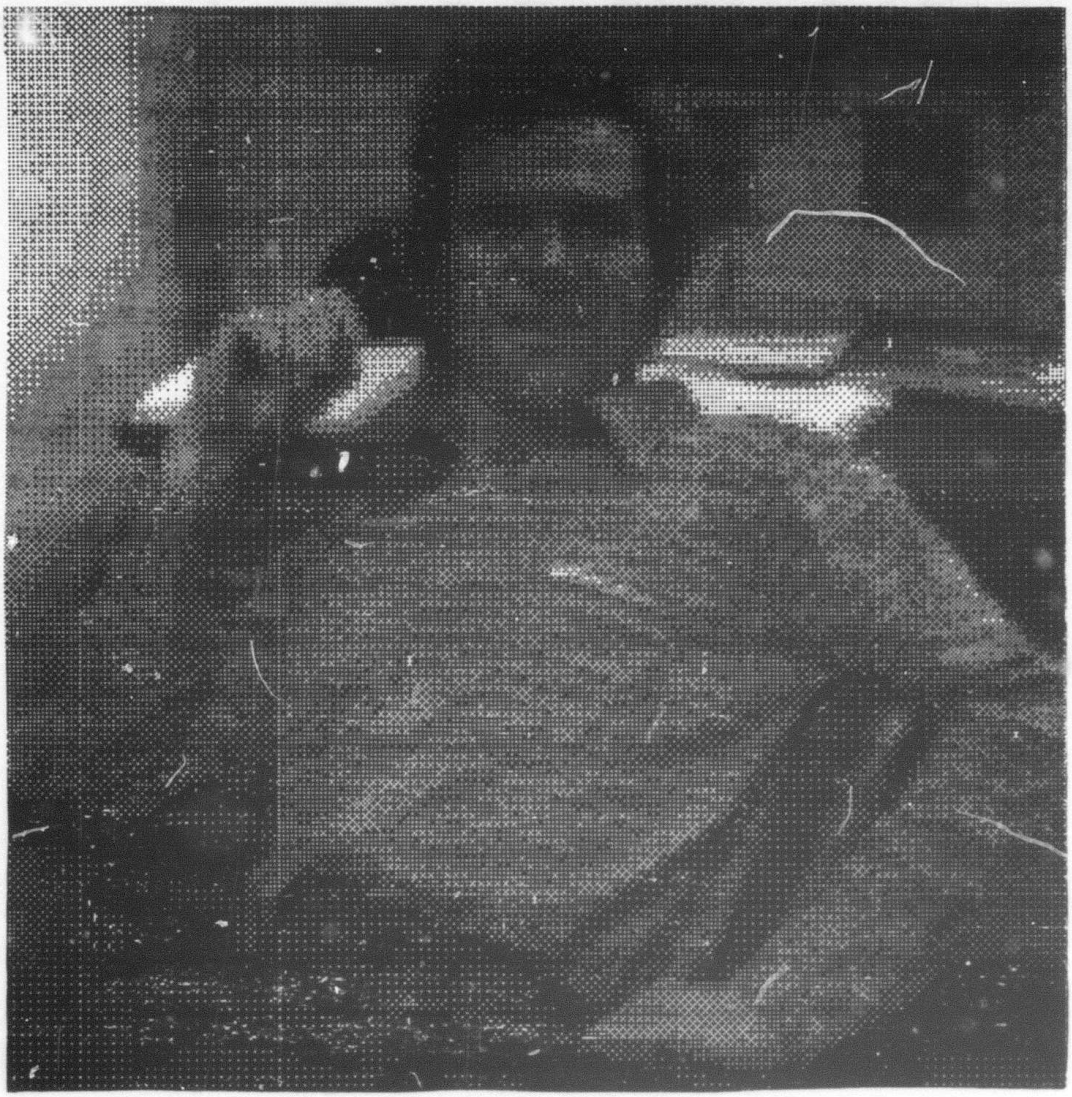


Fig. 1 Original Image

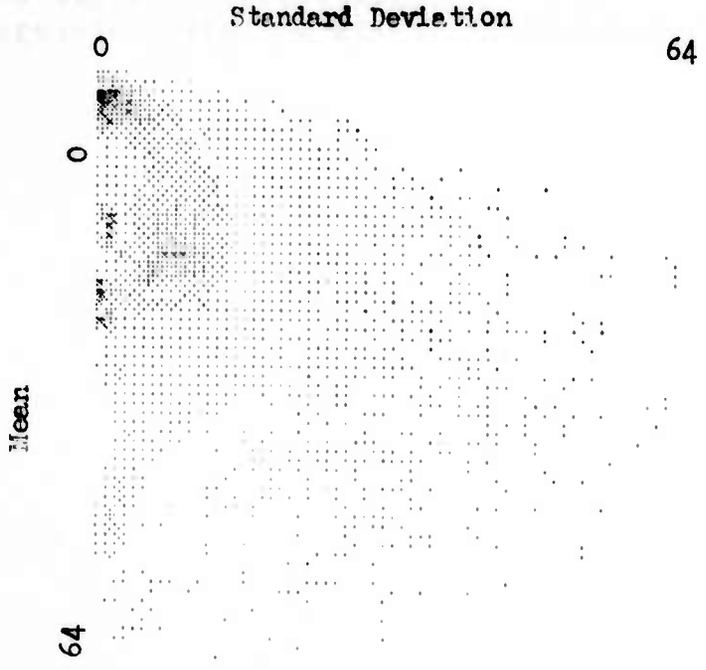


Fig. 2 Mean - Standard deviation feature space

SYMBOL	MINIMUM	MAXIMUM
.	-1.88333E+01	1.88333E+01
•	1.88333E+01	5.65000E+01
-	5.65000E+01	9.41667E+01
■	9.41667E+01	1.31833E+02
■	1.31833E+02	1.69500E+02
■	1.69500E+02	2.07167E+02
■	2.07167E+02	2.44833E+02
■	2.44833E+02	2.82500E+02
■	2.82500E+02	3.20167E+02
■	3.20167E+02	3.57833E+02

Fig. 3 Ranges of values (in picture elements) for symbols in Parts (a) of Figs. 4-6

Fig. 4(a) Feature space

Fig. 4(b) Image space

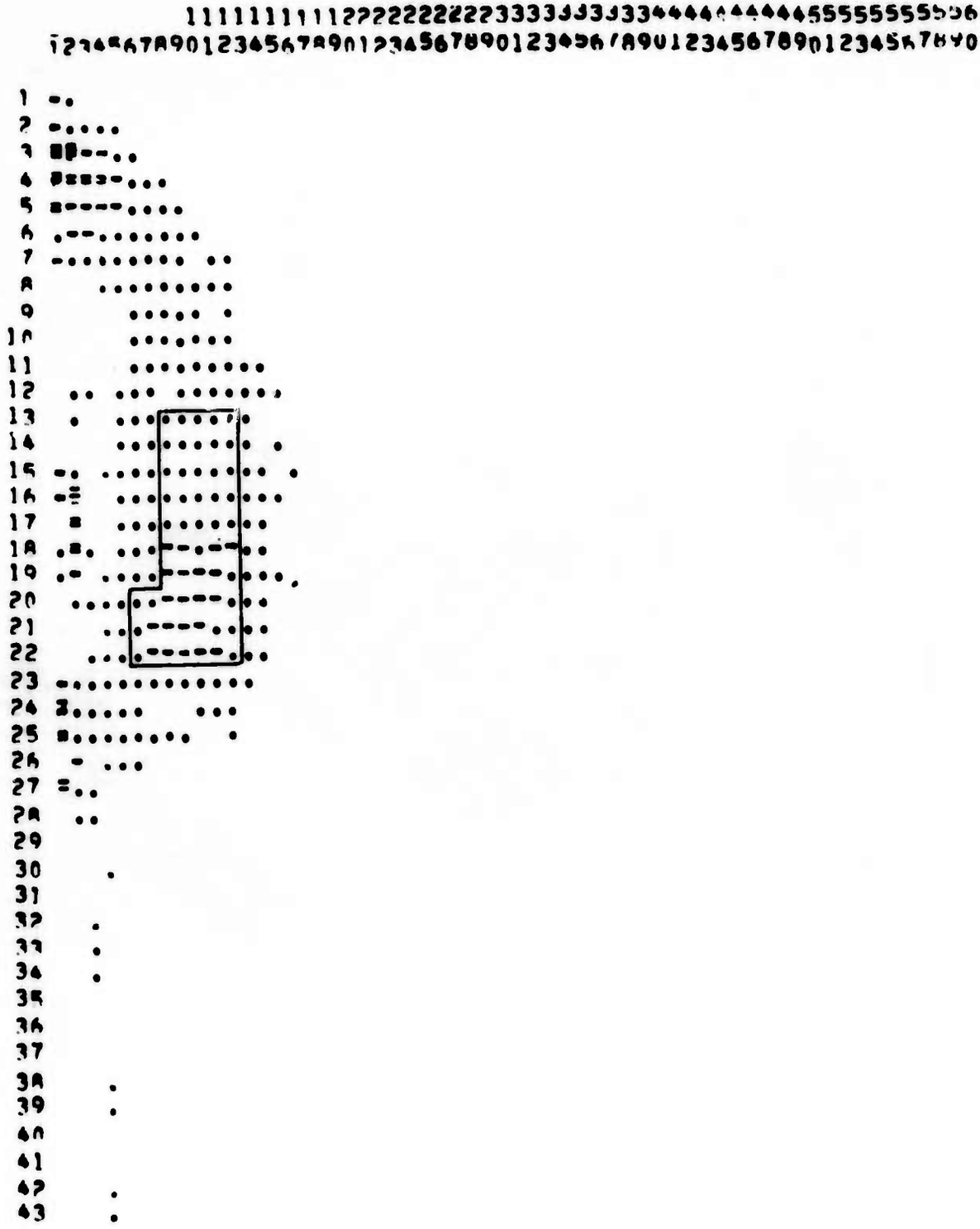


Fig. 5(a) Feature space

Fig. 5(b) Image space

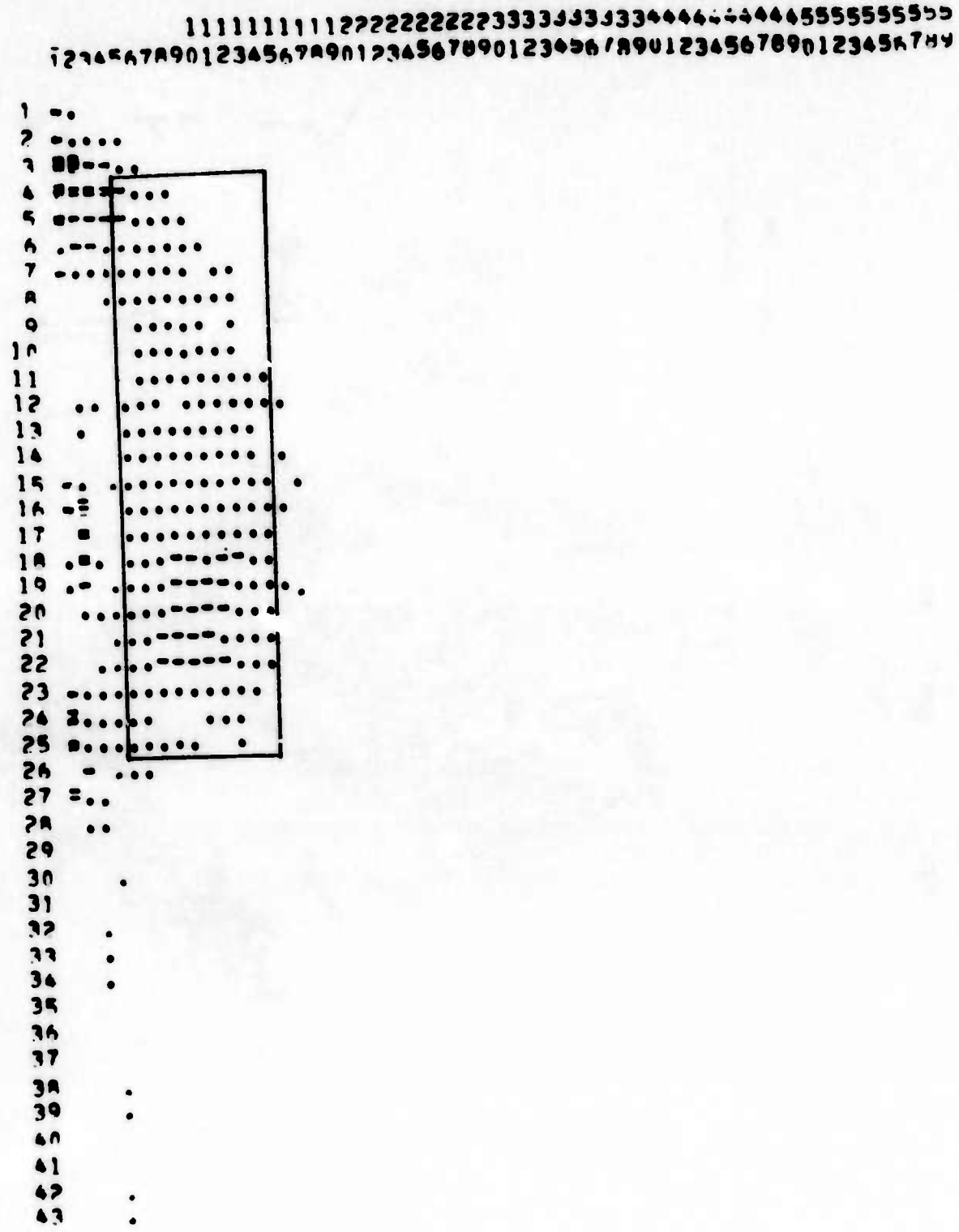


Fig. 6(a) Feature space

Fig. 6(b) Image space

NOISE REDUCTION IN PHOTOGRAPHIC IMAGES

M. Y. Yoo and T. S. Huang

I. INTRODUCTION

Noise reduction using Wiener filters tends to smear the edges in the image. Here we report a technique of noise reduction which retains edge sharpness.

One way of decomposing a picture such that recombination of the components results in the original is "synthetic highs" system. In this system the image is separated into the "lows" (low frequency part) and the "edge part" by the Gaussian low pass filter and the Laplacian (or gradient) edge detector. To recover the original the edge is passed through the reconstruction filter (which is related to the low pass filter) before the recombination with the lows. In this paper we use the synthetic highs system for reducing the noise in the original. Specifically we used additive Gaussian noise and generated noisy picture with S/N = 10 dB. We defined the signal-to-noise ratio of the image as the ratio of the RMS signal to the RMS noise. Low pass filter has a smoothing effect and "lows" does not include much noise. But the output of the edge detector is quite noisy. If there is any way to remove this noise, we can recover a clean picture. The simplest way is to threshold the edge picture. A better way is to detect edges in the noisy picture and then use these edge locations as a mask to eliminate noisy points in the "edge part" of the decomposed picture. So far no really good boundary detection algorithm of noisy picture has been found. One fairly satisfactory way proposed by Tang and Huang [1] (see Tang and Huang in this report) is used for our experiment. Fig. 1 shows the schematic diagram of our experiment.

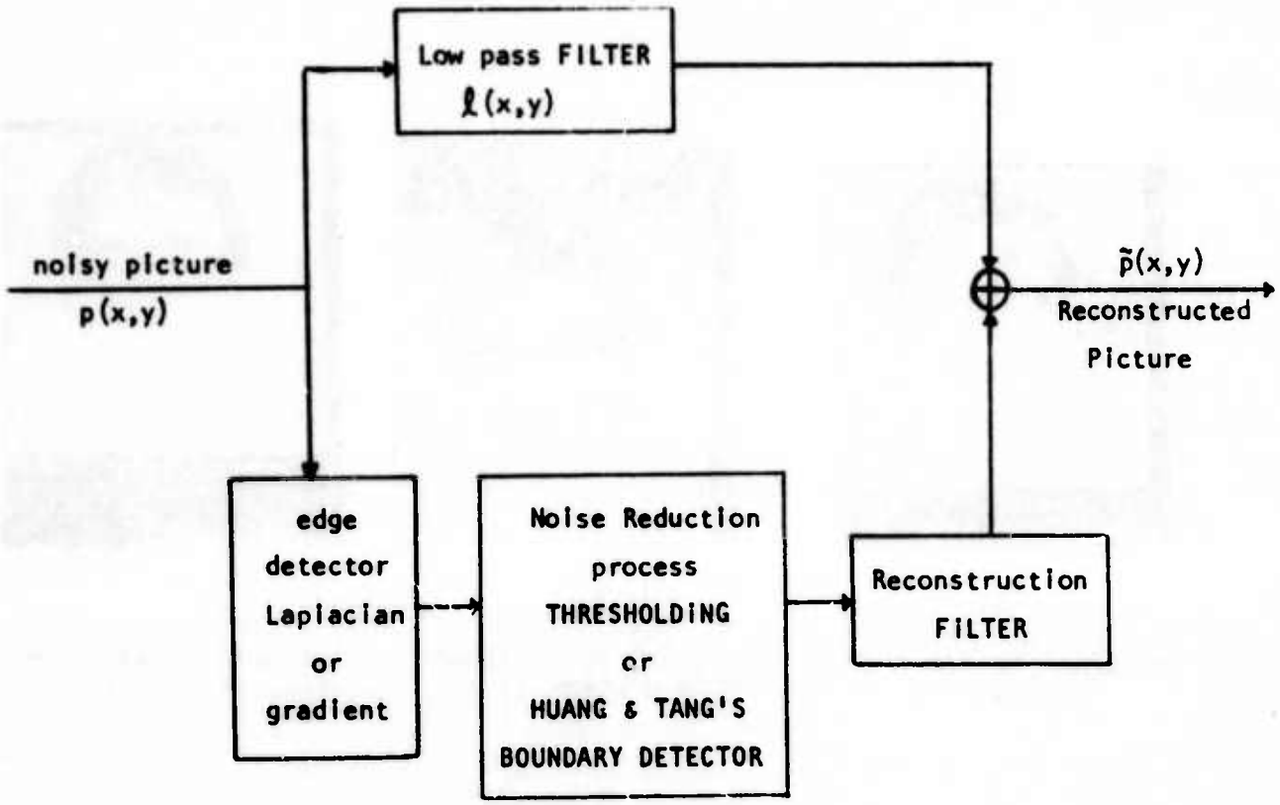


Fig. 1 Block Diagram

II. LOWS

We used a Gaussian low pass filter with impulse response

$$\ell(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2 + y^2}{2\sigma^2}}$$

The frequency response is

$$L(fx,fy) = e^{-2\pi^2\sigma^2 (fx^2 + fy^2)}$$

If we define the bandwidth of the Gaussian filter as the standard deviation of the filter, the bandwidth is

$$B = \frac{1}{2\pi\sigma}$$

(a)
Original Noisless Image

(b)
Original Noisy Image (S/N = 10 dB)

(c)
Averaged Noisy Image (2x2 window used)

Fig. 2

(a)
Variance = 1.03
Bandwidth = 0.155

(b)
Variance = 2.06
Bandwidth = 0.077

Fig. 3 Low passed image

(a)
Threshold = 80
Bandwidth = 0.274

(b)
Threshold = 100
Bandwidth = 0.274

(c)
Threshold = 100
Bandwidth = 0.155

(d)
Threshold = 200
Bandwidth = 0.155

Fig. 4 Reconst. Picture (Thresholding used)

(a)
3x3 Window
Bandwidth = 0.274

(b)
5x5 Window
Bandwidth = 0.155

(c)
5x5 Window
Bandwidth = 0.155

Fig. 5 Reconst. Picture
(Edge detection used).

III. HIGHS

The Laplacian (or gradient) edge detector and reconstruction filter is discussed in Graham [2] and we just describe the noise reduction process here. The simple thresholding is straight forward and needs no discussion.

Tang and Huang's boundary detection algorithm uses the combination of averaging and the gradient detection. Once we get the boundary (from Fig. 2(b)) by this method we may use them as a mask on the Laplacian picture. All points in the Laplacian picture not falling in the boundaries are set to zero. But before applying this mask, there is one thing to worry about. Many of the pictures have ramp function type edges and the Laplacian edge detector gives two spikes (positive and negative) along the boundary. Therefore, the boundaries are expanded to include both spikes. We used 3x3 and 5x5 windows for this purpose: each boundary point is expanded to a 3x3 or 5x5 block.

The reconstructed picture is quite improved compared with the noisy original. We also tried the attenuation of the high part before the recombination with lows, and averaging of the noisy picture. The results are shown in Figs. 2-5.

IV. CONCLUDING REMARKS

We used the Laplacian edge detector for this experiment. But the gradient edge detector may be better than the Laplacian for the noisy picture and we will continue this work with the gradient.

REFERENCES

[1] G. Y. Tang and T. S. Huang, "Edge Extraction Techniques," this Report,

[2] D. N. Graham, "Image Transmission by Two-Dimensional Contour Coding," Proc. IEEE, Vol. 55, No. 3, March 1967.

SEGMENTING COLOR PICTURES WITH BLOB

Paul A. Wintz

The standard color pictures shown in Fig. 1 was first transformed from the red, green and blue coordinate system through Y, i, and Q coordinate system. The BLOB algorithm was then run on the Y, i, Q pictures. All of the picture elements in each BLOB in the Y, I, Q domain were replaced with an average value. The inverse transform from the Y, i, Q domain back to the red, green and blue domain was then made. The resulting pictures for several different settings of the thresholds are represented in Figs. 2, 3, and 4. The reconstructed pictures in Figs. 2, 3, and 4 show only a small amount of distortion from the original indicating the possibility of using BLOB as a coding algorithm since significant compression ratios can be obtained by coding the BLOB's relative to coding the original picture elements

Reference

1. P. A. Wintz and J. N. Gupta, "A Boundary Finding Algorithm and its Applications," IEEE Trans. on Circuits and Systems, March 1975.

Fig. 1 Original image

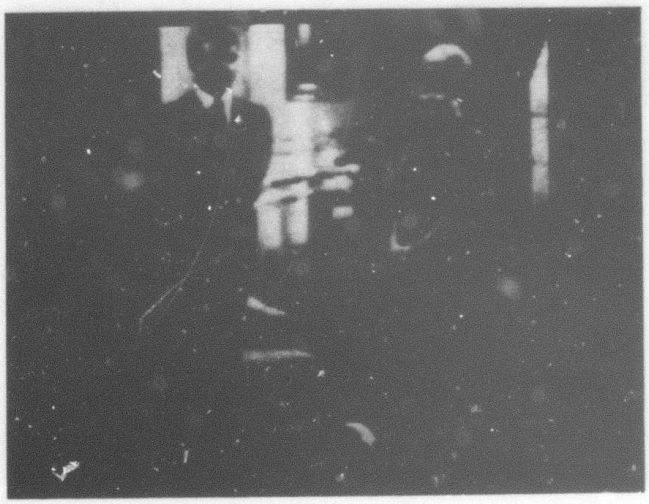


Fig. 2 Coded image .78 bit/picture element

Fig. 3 Coded image .65 bit/picture element

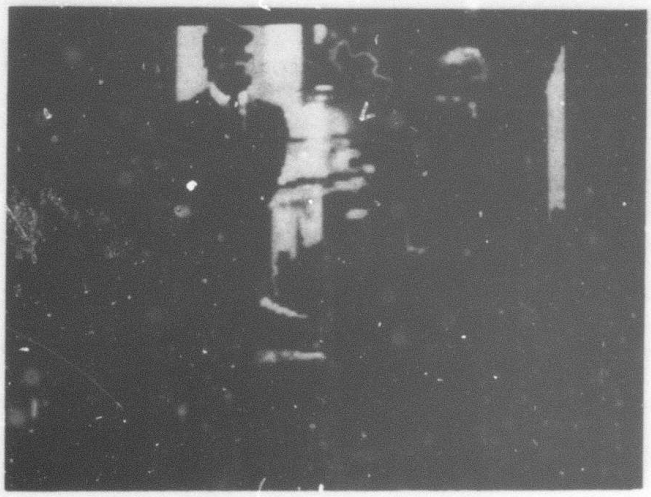


Fig. 4 Coded image .45 bit/picture element

STATISTICAL MEASURES FOR THE CLASSIFICATION OF TEXTURES

O. R. Mitchell and C. R. Myers

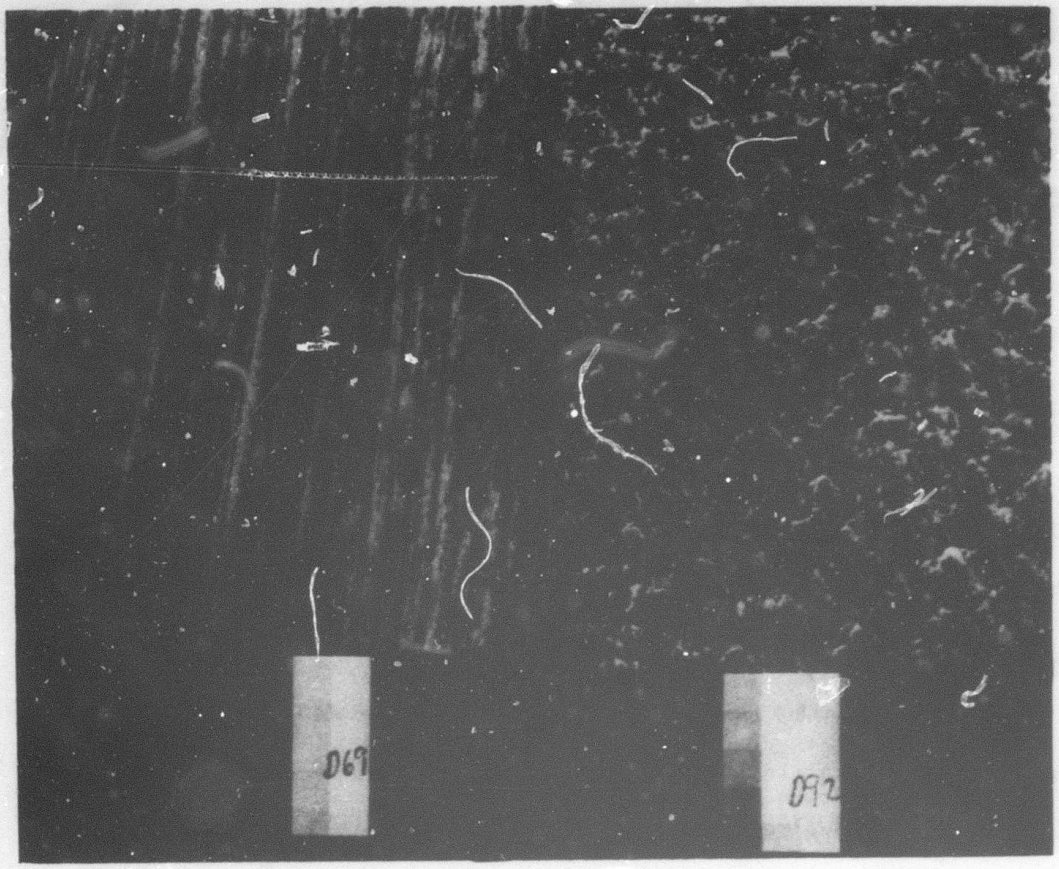
The need to classify image regions has led us to search for optimum statistical measures (or features) for the classification of image textures. First, we have sought to improve existing measures and second, we have derived promising new measures.

The statistical measures of texture based on spatial dependence probabilities (or the neighboring grey tone co-occurence properties) has been one of the most promising [1]. However, these measures are not invariant to resolution and most of them are not invariant to illumination changes.

BRIGHTNESS INVARIANCE

The effect of a change in illumination level (including shade vs sun and film development differences) is primarily a multiplicative effect. To make our texture measures invariant to this effect we have taken the logarithm of all intensities and subtracted the local mean. This is a form of homomorphic filtering and is essentially equivalent to the first processing stages in the human visual system (a logarithmic transformation followed by a band-pass filter) [2].

For initial tests we have used texture samples from a book by Brodatz [3], Figure 1 shows typical textures. We used ten different textures, 5 samples of each for training and 2 samples of each for testing. Each sample was 64x64 points, with 64 grey levels. The features used for classification were those suggested by Haralick [2]. The classification procedure was a simple normalized Euclidian distance measure with all features weighted equally. The results are shown in Fig. 2. The logarithmic measure gave results comparable to the standard method with these textures. However, the



Wood Pigskin

Fig. 1 Texture Samples

10 Textures
5 Training of each
2 Test of each

70 Total classified texture areas
64 x 64, 64 levels

Spatial Dependence, Distance = 1

	Training % accurate	Test % accurate
Normal	84	60
Log	82	80
Log, subtract mean	84	80
Hist. equalization	36	30

Fig. 2 Texture Classification Results

logarithmic measure is invariant to illumination changes which would deteriorate the results of the standard method considerably.

RESOLUTION INVARIANCE

In order to make the spatial dependence matrix measures of texture invariant to resolution (includes camera or scanner distances and sampling rate) we have used the two-dimensional regional autocorrelation to regulate the distances used in calculating the spatial dependence matrix. Typical autocorrelations (actually autocovariances) are shown in Figure 3. We have used the contour at which the autocorrelation has fallen to a fixed percentage of its maximum value as the primitive texture shape. This contour is then used for selecting horizontal and vertical distances to be used in calculating the spatial dependence matrix. Results using the same texture training and test samples as before are shown in Figure 4. When a contour of 20% maximum value is used the results are very encouraging. This measure is invariant to changes in resolution and sampling rate (as long as the image is not undersampled).

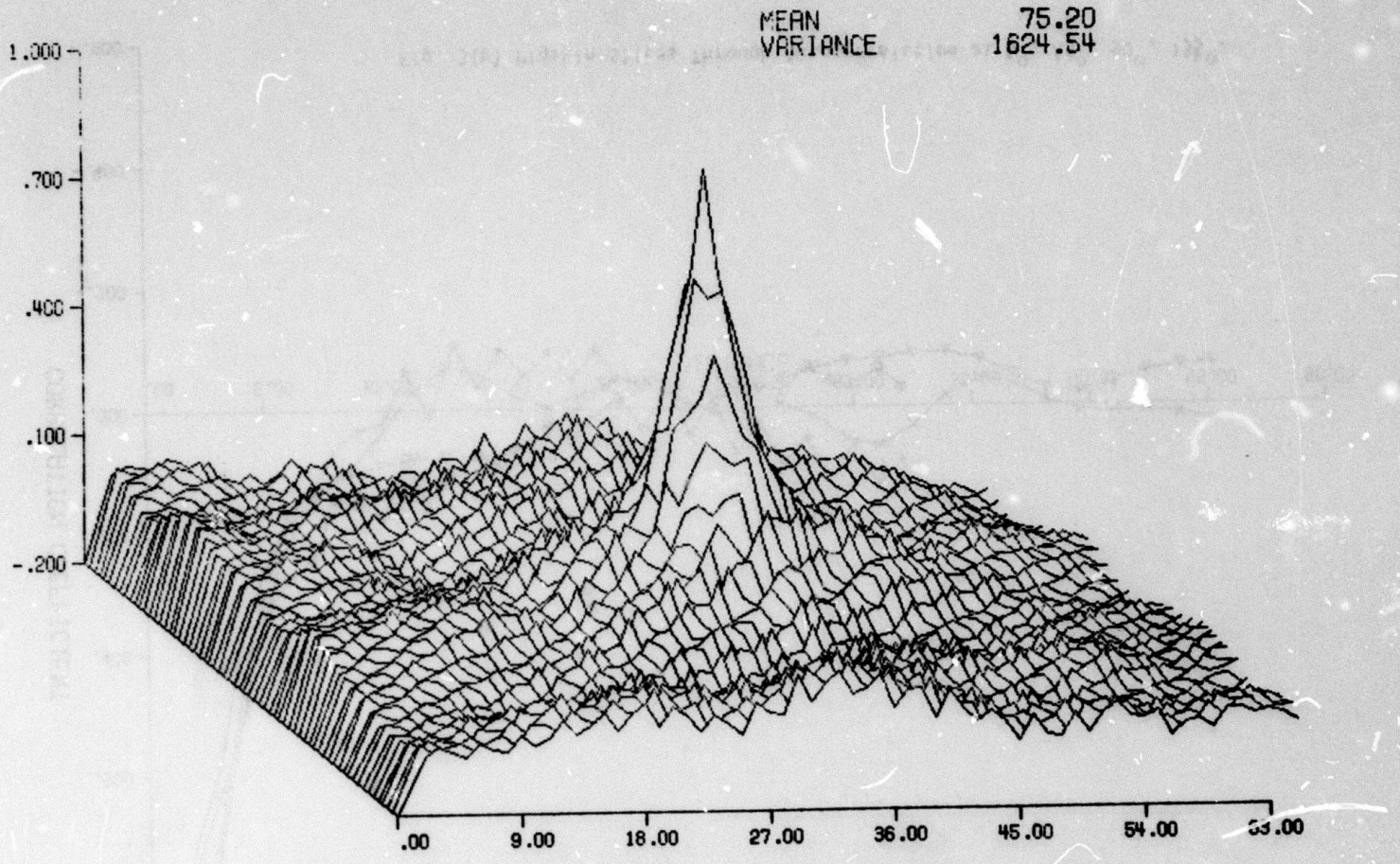


Fig. 3(a) Pigskin Autocorrelation

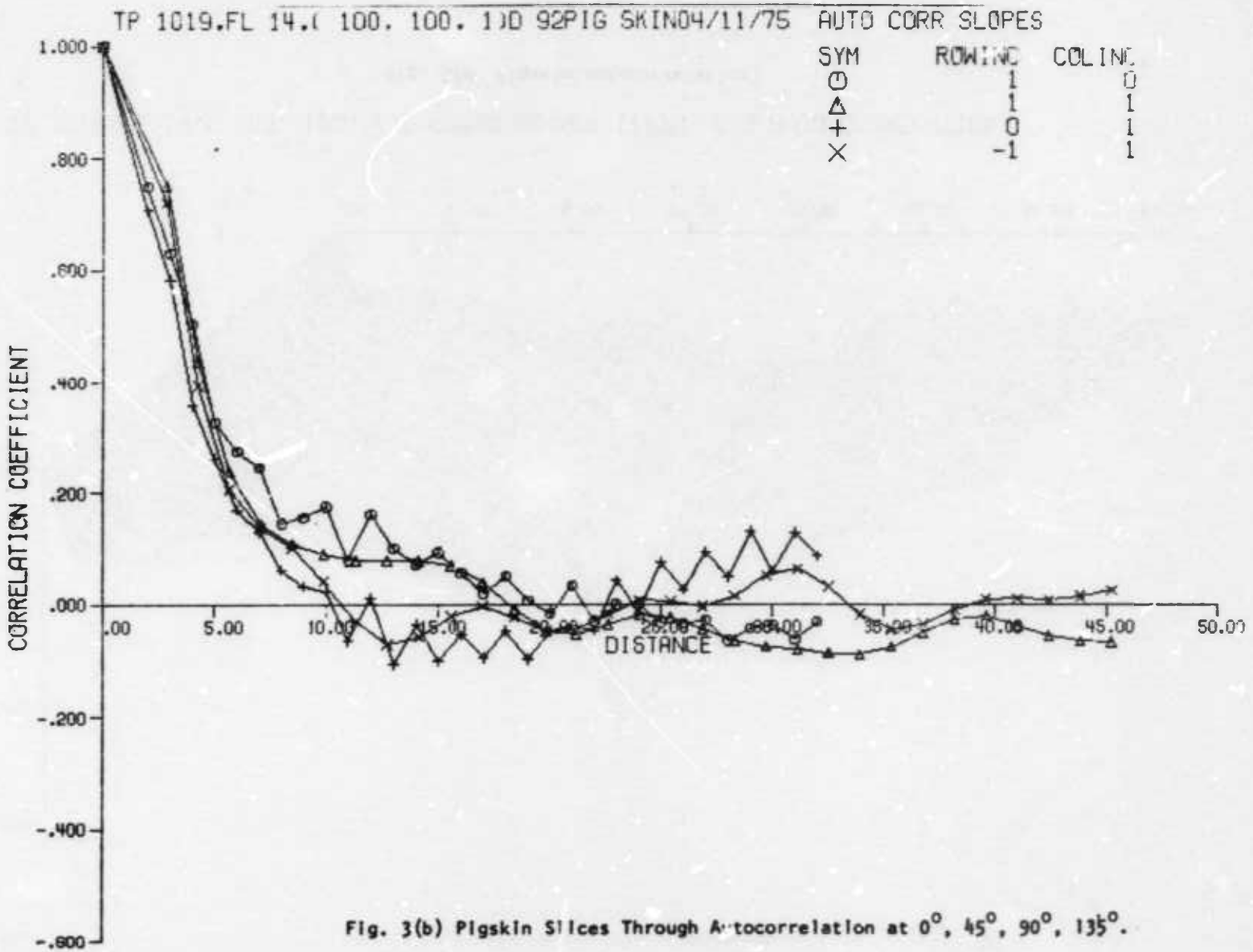
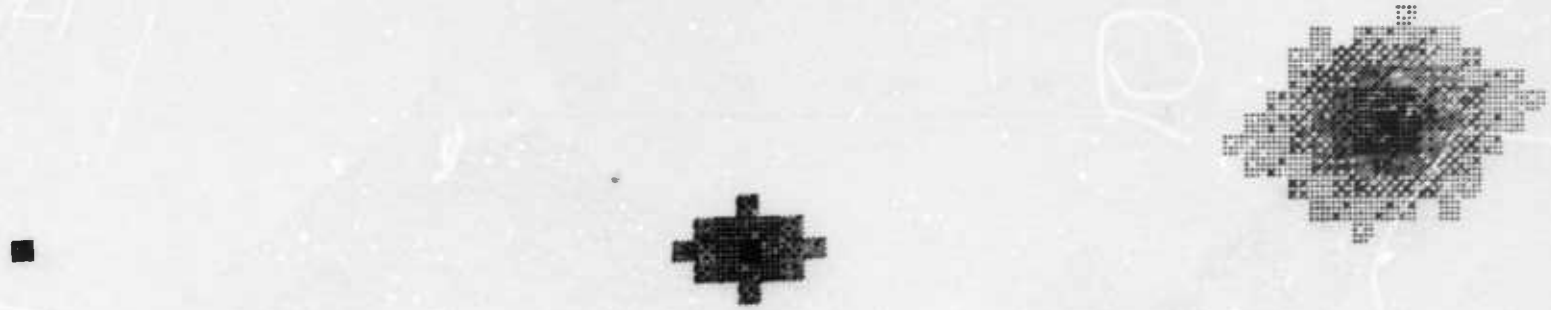


Fig. 3(b) Pigskin Slices Through Autocorrelation at 0°, 45°, 90°, 135°.



80% Correlated

50% Correlated

20% Correlated

Fig. 3(c) Pigskin Threshold Contours

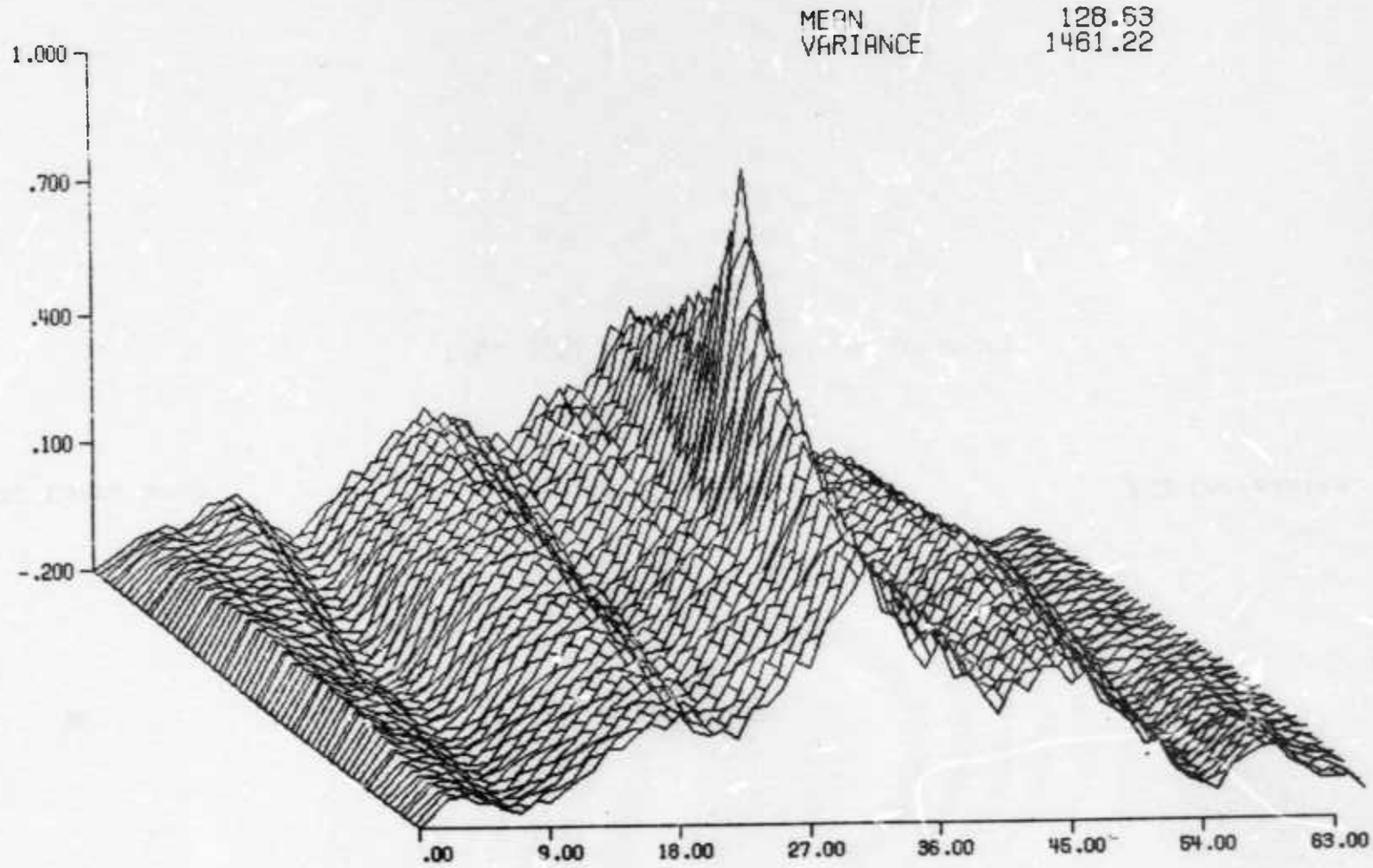


Fig. 3(d) Wood Autocorrelation

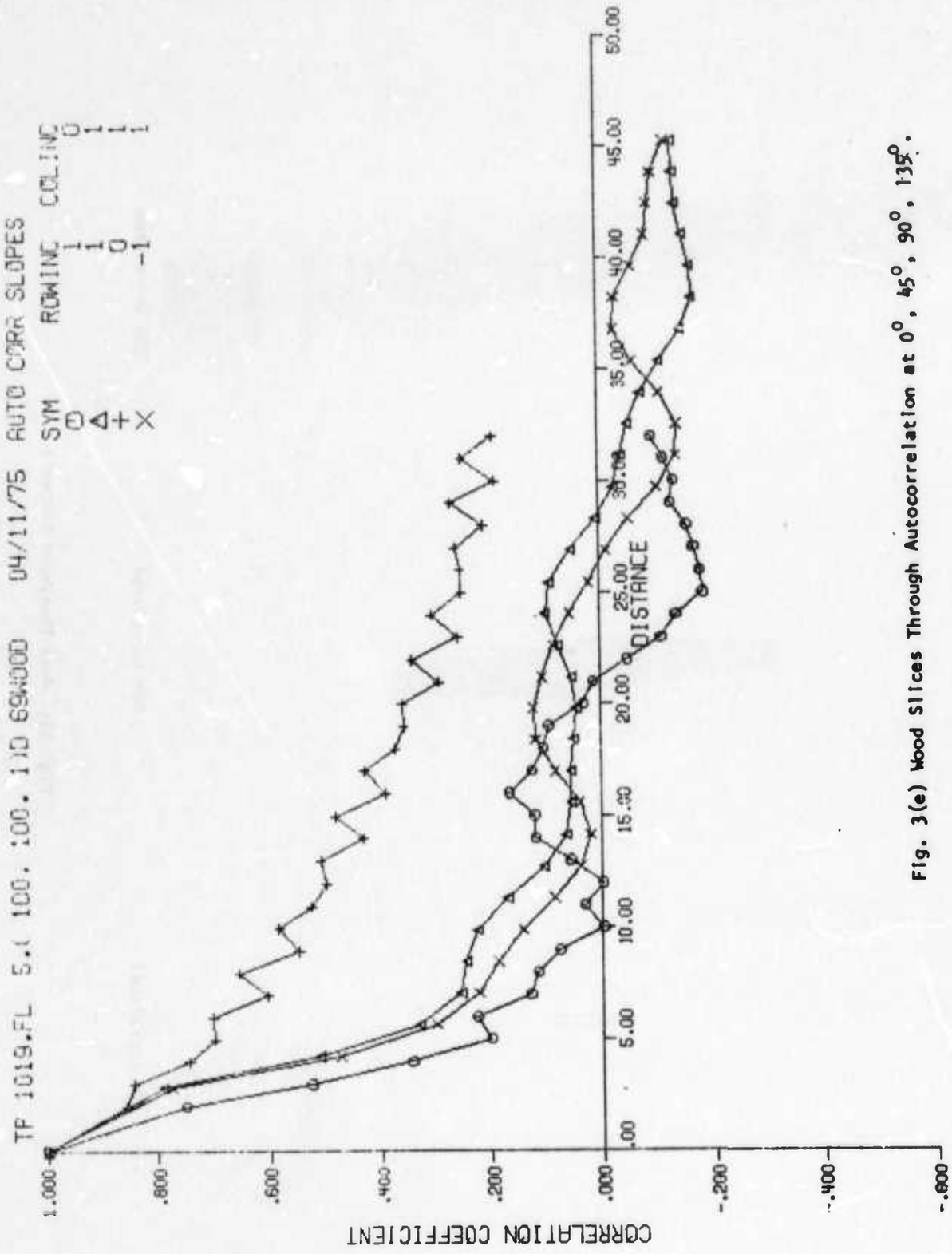
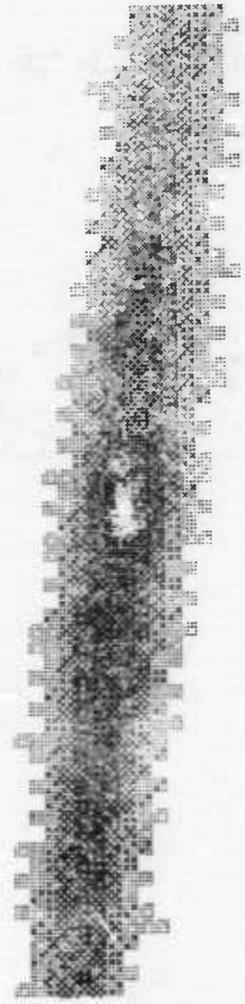


Fig. 3(e) Wood Slices Through Autocorrelation at 0°, 45°, 90°, 135°.

80% Correlated

50% Correlated



20% Correlated

Fig. 3(f) Wood Threshold Contours

Distance depends on autocoovariance, logged intensities

	Training	Test
c = .8	79%	70%
c = .5	83%	70%
c = .2	95%	80%

Fig. 4 Variable Distance Spatial Dependence Classification Results

A NEW STATISTICAL MEASURE

A new measure for texture classification has been developed which is computationally much simpler than the above methods. It is based on the human visual system intuition that the important texture information is contained in the relative frequency of local extremes of various sizes in intensity. The principal measure in this process is the determination of the number of maxima and minima, along a one-dimensional scan direction. The extremes are not included unless they meet a threshold condition as shown in Figure 5. The maximum is called a maximum only if the intensity falls the threshold amount below the maximum before a higher valued intensity is encountered. This hysteresis effect ignores small ripples while including the larger ones. By repeating the process for several threshold settings, a vector of numbers is obtained which characterizes the texture.

We have found that the steps necessary to make this measure invariant to illumination and resolution are very easily implementable and improve the accuracy of the results. These steps are (1) take the logarithm of the intensities first; (2) use the ratios of the number of extrema at each threshold to the next instead of numbers of extrema themselves. The threshold levels used for these measures are shown in Figure 6. The results are shown in Figure 7. We used 7 features in both horizontal and vertical directions for a total of 14 features. The results are very encouraging indeed.

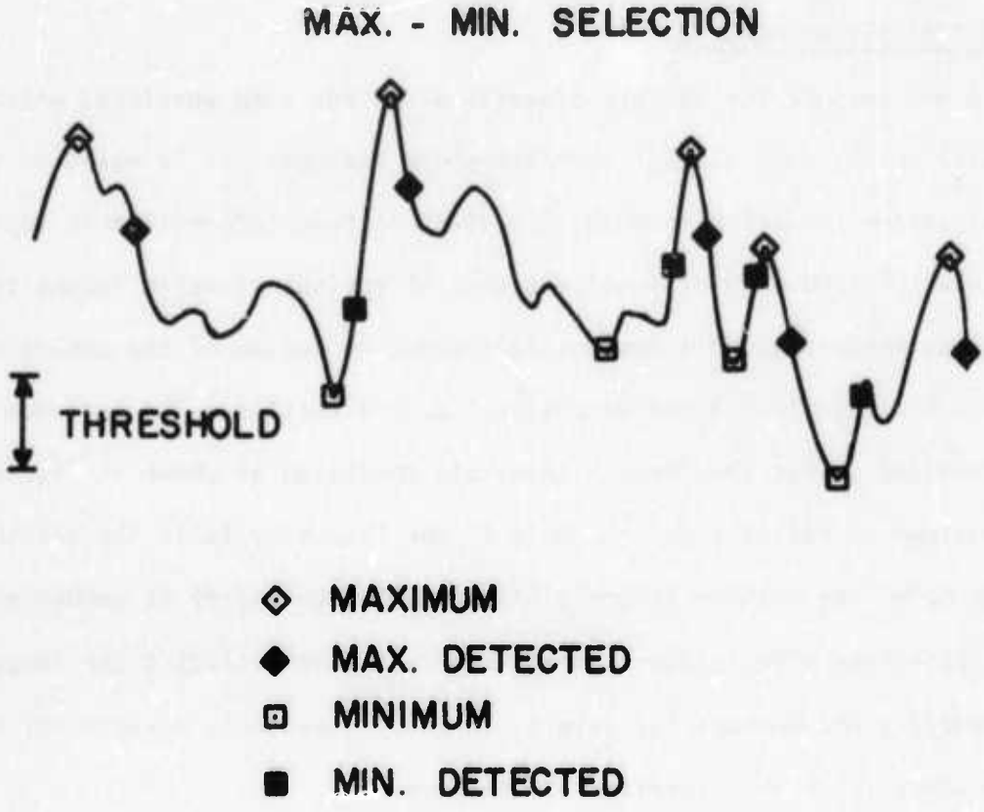


Fig. 5 MAX-MIN Selection

a	a/b	b-c
b	b/c	c-b
c	c/d	d-c
d	d/e	e-d
e	e/f	f-e
f	f/g	g-f
g	g	g
Counts	Ratios	Difference

Where threshold is

a = 128		a = log 128
b = 64		b = log 64
c = 32		c = log 32
d = 16	for normal and	d = log 16 for logarithmic
e = 8		e = log 8
f = 4		f = log 4
g = 2		g = log 2

Fig. 6 MAX-MIN Features Studied

Max - Min Method	Training % correct	Test % correct
Counts, normal	82	40
log	72	80
Difference, normal	90	77
log	94	80
Ratios, normal	94	70
log	94	80
Mean and variance, normal	54	20
log	80	60

Fig. 7 Results MAX-MIN Method

CONCLUSION

We have derived feature classification techniques which are invariant to illumination and resolution changes. A new technique is described which allows simple computation of texture features. Although our sample sizes are small, we feel the results do show that the techniques described can be optimized to provide large improvements in texture classification techniques.

REFERENCES

[1] R. M. Haralick, K. Shanmugan, and I. Dinstein, "Textural Features for Image Classification," IEEE Trans. on Systems, Man, and Cybernetics, Vol. SMC-3, No. 6, pp. 610-621, November 1973.

[2] T. G. Stockham, "Image Processing in the Context of a Visual Model," Proceedings of the IEEE, Vol. 60, No. 1, pp. 828-842, July 1972.

[3] P. Brodatz, Textures, Dover Publications, New York, 1966.

AN OPTIMAL FEATURE SUBSET SELECTION ALGORITHM

K. Fukunaga and P. M. Narendra

Patterns to be classified by a recognition system are generally characterized by a set of measurements or features. Often, the dimensionality of this feature space is too large for efficient and reliable application of existing classification techniques. The feature extraction problem is then, to reduce the dimensionality of the feature space, without significantly affecting the discriminatory capacity of the feature set. Classification with the reduced set of features is then facilitated.

One approach to feature extraction is to select a smaller subset of m features of the set of n original features ($m < n$). There are $\binom{n}{m} = \frac{n!}{m!(n-m)!}$ such subsets of size m. The best subset of size m is then the subset for which the value of a criterion is extremized over all subsets of size m. One way to obtain the best subset is of course, to evaluate the criterion for all the $\binom{n}{m}$ subsets and choose the one that yields the extreme value. But exhaustive search is computationally unfeasible even for modest values of n and m, as the amount of search involved becomes prohibitive. For example, with n=24 and m=12, the number of subsets to be considered is 2,704,156. Several approaches have been suggested to select a good subset without exhaustive enumeration. Forward selection [1], and Dynamic Programming techniques [2,3] are among these. But none of these techniques guarantee that the selected subset is indeed the best among all subsets of size m.

We have formulated the feature subset selection problem as a combinational optimization problem. A branch and bound search algorithm [4,5] has been developed to find the best subset without exhaustive search. The

algorithm is efficient and is guaranteed to be optimal. The algorithm involves recursive enumeration of the subsets. Unfeasible subsets are discarded without being explicitly enumerated. The algorithm is stated for any general criterion; recursive equations are developed to aid the implementation of the algorithm for a class of quadratic criteria such as the Fisher criterion, Divergence and the Bhattacharyya Distance.

To test the effectiveness of the feature selection algorithm, it was applied to multispectral data from airborne scanners. The data originated from the Laboratory for Applications of Remote Sensing (LARS) and consisted of classified data points of fields planted with corn, soybean, rye, wheat, etc. The problem was to select out of the 12 data channels encompassing the entire spectrum (0.4 to 1.0 micrometer) the subset of 4 channels that yields the optimum values of a criterion. The Fisher criterion was chosen. Exhaustive search involves the computation of all the $\binom{12}{4} = 495$ subsets of channels to obtain the best subset. The present algorithm was applied to obtain the best subset to discriminate soybean from corn. The computational effort was reduced by a factor of more than 50 over exhaustive enumeration, at the same time guaranteeing that the chosen best subset was indeed the best among all the 495 subsets of size 4.

Further, the algorithm was applied to choose the best subset of 12 out of 24 channels. There are 2,704,150 candidate solutions. The algorithm obtained the best subset with the computational effort equivalent to computing the criterion for 6000 subsets. The savings are indeed substantial.

This work is being documented for publication. Further work is underway to extend the approach to other problems in Pattern Recognition.

REFERENCES

[1] A. N. Mucciardi and E. E. Gose, "A Comparison of Seven Techniques for Choosing Subsets of Pattern Recognition Properties," IEEE Trans. Comput. Vol. C-20, pp. 1023-1031, September 1971.

[2] C. Y. Chang, "Dynamic Programming as Applied to Feature Subset Selection in a Pattern Recognition System," IEEE Trans. Syst, Man, Cybern, Vol. SMC-3, pp. 166-171, March 1973.

[3] G. D. Nelson and D. M. Levy, "A Dynamic Programming Approach to Selection of Pattern Features," IEEE Trans. Syst., Sci., Cybern, Vol. SSC-4, pp. 145-151, July 1968.

[4] S. W. Golomb and L. D. Baumert. "Backtrack Programming," Journ. ACM, vol. 12, pp. 516-524, January 1965.

[5] E. L. Lawler and D. E. Wood, "Branch-and-Bound Methods - A Survey," Oper. Res. vol. 149, No. 4, 1966.

EXTENSION OF THE SUBSPACE METHOD IN PATTERN RECOGNITION TO PICTURES

G. V. Sherman and K. Fukunaga

I. INTRODUCTION

Most decision-theoretic pattern recognition schemes are of the zonal variety. That is, observations are made in an n-dimensional vector space which has been partitioned into M zones, corresponding to the M classes, by (n-1)-dimensional surfaces. Fig. 1 illustrates this concept for n=2 and M=3. Ideally the partitions in Fig. 1 are determined by Bayes Law. However, usually the n-dimensional class dependent densities are unavailable and too cumbersome to compute. Hence, suboptimum methods are used to determine the partitions.

Watanabe [1] has proposed a different approach. Instead of partitioning the observation space into zones, perhaps one should find feature subspaces which correspond in a one-to-one manner with the classes. Fig. 2 illustrates this concept for n=3 and M=2.

Wantanabe's original approach can be described as follows:

1. Construct a representation subspace for each class from the eigenvectors of the class autocorrelation matrices; i.e., the discrete Karhunen-Loeve, (K-L) expansion. Retain only those eigenvectors, corresponding to the largest eigenvalues, which contain t% of the sample energy on the average. Typically $t \simeq 95$ and is the same for each class.
2. Construct a feature subspace for each class by subtracting from the corresponding representation space its intersections with the representation spaces of other classes. The feature space are then transformed so that they are all mutually orthogonal.
3. That portion of the observation space which is not included in any feature space makes up the reject space.

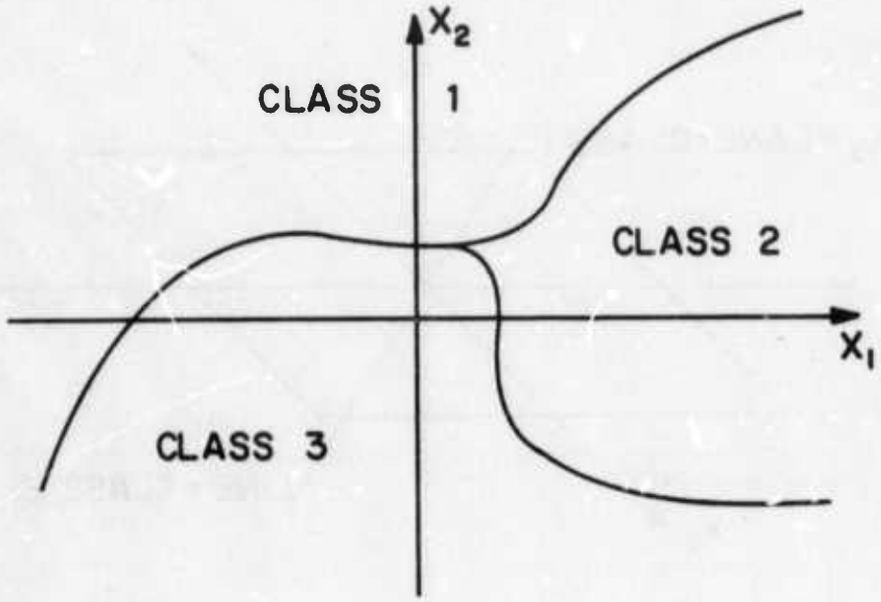


Fig. 1 Two-Dimensional Observation Space Partitioned for Three Classes

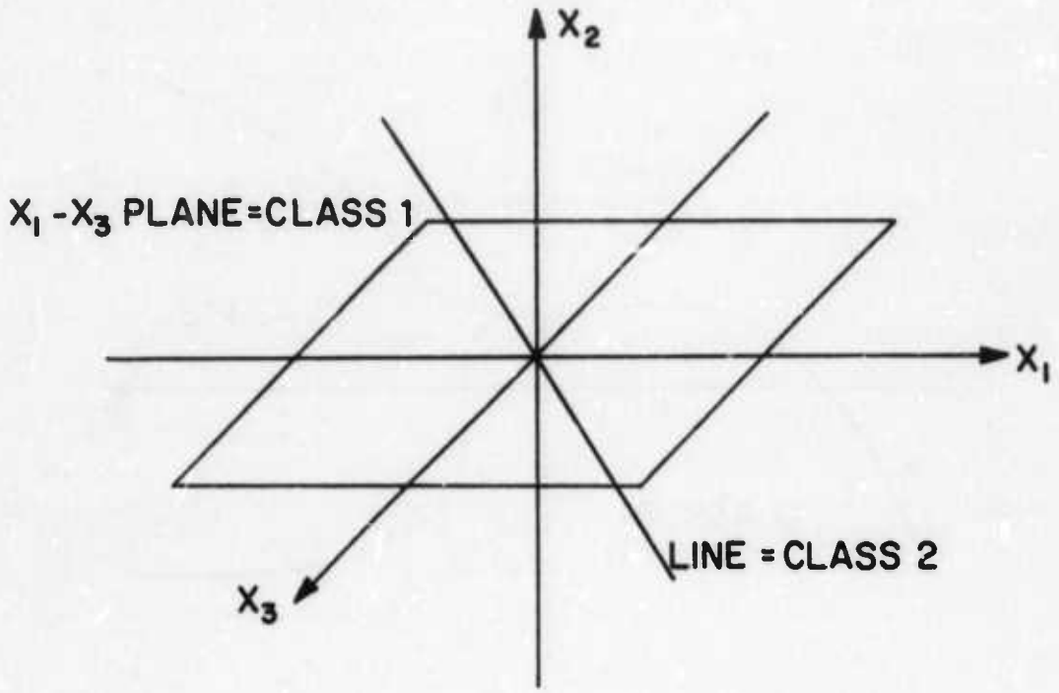


Fig. 2 Three-Dimensional Observation Space with Two Feature Subspaces

4. Unknown samples are projected onto the various class feature spaces and classification is accomplished by a Bayes rule which is based on membership in a subspace rather than membership in a zone.

Therrien [2] made the subspace method more practical by deriving efficient method for computing the intersection and union of subspace.

II. EXTENSION TO PICTURES

The representation spaces for different classes of pictures cannot be obtained by the Karhunen-Loeve approach. The dimensionalities involved are too high. Therefore, the pictures are represented by an outer-product expansion of the form

$$A = \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_{ij} \phi_i \psi_j^T \qquad (1)$$

where A is an nxn raster-scanned picture, γ_{ij} is a scalar; and ϕ_i, ψ_j are n-dimensional vectors. The $\{\phi_i\}$ are $\{\psi_j\}$ are complete orthonormal bases which span the column and row spaces of A respectively. That is, any column of A can be expressed as a linear combination of ϕ_i's and similarly for rows of A and ψ_i's. The $\{\phi_i\}$ and $\{\psi_j\}$ bases can be obtained by many methods, fixed bases such as Fourier and Hadamard, or data dependent bases such as discrete K-L. The latter was choosen for our experiments. Note than an n^2-dimensional K-L expansion has been replaced by two n-dimensional K-L expansions.

Furthermore, we have shown that the intersection and union of outer-product subspaces can be calculated by finding the intersection and union of the corresponding column and row spaces separately. Again the result is that an n^2-dimensional eigenvector calculation has been replaced by two n-dimensional eigenvector calculations. This result does not depend on the choice of $\{\phi_i\}$ and $\{\psi_j\}$.

Feature extraction is completed by transforming the feature spaces so that they are all mutually orthogonal. This operation originally would require an n^2-dimensional matrix inversion. But we have shown it to be equivalent to two n-dimensional matrix inversions operating again on the row and column spaces separately.

The subspace classifier proposed by Watanabe [1] was based on heuristic arguments. It requires the calculation of the following quantities:

1. $p(\ell|\omega_k)$, the probability of a class-k sample "appearing" in subspace ℓ.
2. $p(\omega_k)$, the a priori probability of class k.
3. $q(\omega_k|\ell)$, the probability of a sample being in class ω_k given that it "appears" in subspace ℓ.
4. $W(\ell|\underline{X})$, the probability of random sample $\underline{X}$ appearing in subspace ℓ.
5. $q(\omega_k|\underline{X})$, the probability of $\underline{X}$ being from class ω_k.

A serious question arises concerning this classifier. Can we define the event "$\underline{X}$ appears in subspace ℓ"? In most cases, unknown sample $\underline{X}$ will not be completely contained in any feature space. However, we intuitively feel that $\underline{X}$ appears in subspace S_1 to a greater extent than it appears in subspace S_2 if the projection of $\underline{X}$ on S_1 has a larger norm than that of the projection of $\underline{X}$ on S_2. Thus "$\underline{X}$ appears in subspace ℓ" is a fuzzy event [3,4], Zadeh shows how to define probability of a fuzzy event and also defines conditional probability and Bayes rule for fuzzy events. This formalism results in a different definition for the empirical calculation of $p(\ell|\omega_k)$. All other quantities remain the same.

III. EXPERIMENTAL RESULTS

The pictures used for verification of the above theory were the numerals from the Munson alphanumeric data available through Stanford Research Institute. Since we are not interested in implementing a

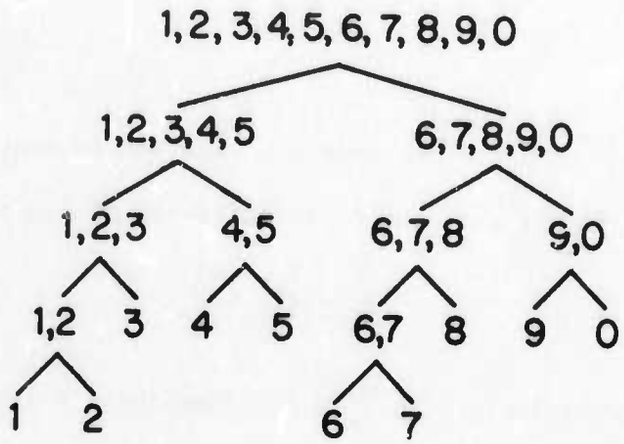


Fig. 3 Layered Machine

character reader, only the numerals were used. Each numeral is a 24x24 binary matrix. 147 samples of each numeral are available. 98 samples were used for design and 49 for testing.

Wantanabe's method for finding feature subspaces was found to be computationally unfeasible, and Therrien's method had a theoretical flaw. Therrien proposed that the ith feature space S_i' should be obtained from the representation space S_k by

$$S_i' = S_i \left(\bigcap_{\substack{k=1 \\ k \neq i}}^{M} \overline{S}_k \right) \tag{2}$$

where $\overline{S}_k$ is the complement of S_k. This approach resulted in empty feature spaces because subspaces do not obey the distributed law of logic given by

$$S_1 \cap (S_2 \cup S_3) \neq (S_1 \cap S_2) \cup (S_1 \cap S_3) \tag{3}$$

That is $S_i \cap S_k = 0$ does not imply $S_i \cap \overline{S}_k \neq 0$. Therefore, a layered machine was implemented in which feature subspaces are formed by

$$S_1' = S_1 - S_1 \cap S_2 \tag{4}$$

That is, the problem was reformulated into the form shown in Figure 3. For our 10-class problem, the layered approach reduced the number of intersection calculations from 90 to 9. Furthermore, (4) resulted in non-empty feature spaces. The classifier itself is yet to be implemented.

IV. FUTURE WORK

As mentioned above, the features must be evaluated by a subspace classifier. Moreover, it is believed that these experiments can reveal the relationship between observation space dimensionality, number of classes,

and number of training samples necessary for valid classifier design and valid feature extraction.

REFERENCES

[1] S. Wantanabe and N. Pakvasa, "Subspace Method in Pattern Recognition," Proceedings on the First International Joint Conference on Pattern Recognition, Washington D. C., pp. 25-32.

[2] C. W. Therrien, "Eigenvalue Properties of Projection Operators and Their Application to the Subspace Method of Feature Extraction," To be published in IEEE Transactions on Computers.

[3] L. A. Zadeh, "Fuzzy Sets," Information and Control, pp. 338-353, June 1965.

[4] L. A. Zadeh, "Probability Measures of Fuzzy Events," Journal of Mathematical Analysis and Applications, pp. 421-427, August 1968.

FOURIER DESCRIPTORS

Paul Wintz and Lyle Stanfield

Shape is an important feature for pattern recognition algorithms. One promising method of shape description is to code a contour (outline of a shape) as a sequence of complex numbers and to work with the Fourier transform of this sequence.

The descriptors used in this project are identical to those described by Granlund [1]. A contour is plotted in the complex plane. As the contour is traced a complex function of arc length is obtained from the coordinates of the points on the contour. The function is periodic since it is possible to trace around the contour more than once. The function is normalized so that its period (or the length of the contour) is 2π. The Fourier transform is taken and is known as the Fourier descriptor of the shape.

The shapes studied in this project were typed on cards and read into a 50x50 array in a FORTRAN program. A routine was written to find an initial point and to trace the contour. The coordinates of the points on the contour were stored in a complex array whose length was adjusted to be a power of two. The adjustment was required by the Fourier transform routine and was accomplished by assuming that points on the contour were connected by straight line and interpolating between points. The interpolation caused some rounding of the corners of objects.

The magnitude and phase of the Foruier coefficients were plotted on the Gould plotter. The magnitude plots are log-log. The a_0 coefficients are shown as *'s on the y-axes.

A routine was also written to set all but certain coefficients to zero. The inverse transforms were then taken to identify the shapes causing certain 'envelopes' in the coefficient plots.

The shapes checked were square, cross, circle, and triangle. The circle and triangle were distorted because of the coarse 50x50 array size. The initial point of all contours was the left-most point at the bottom of the contour. Contours were traced in a counter-clockwise direction.

Fig. 1 shows the square as it was before taking the transform. Fig. 2 shows the magnitude and phase of the Fourier descriptor. On the magnitude plot note that there are four distinct envelopes. Every fourth point seems to be related. Also note that the coefficients a_n where $n = 1 + 4p$ are much larger than the other coefficients, as required by Granlund's symmetry condition.

Figs. 3-6 show reconstructions of parts of the square. Fig. 3 is the inverse transform of the coefficients a_1, a_5, a_9, ..., a_{1+4k}. All other coefficients were set to zero. Due to the square's symmetry these coefficients should contain all of the information about the square. Notice that all of the corners are rounded identically so the contour has a greater amount of symmetry than the original square.

Figs. 4-6 were generated by the sets of coefficients of the forms a_{2+4k}, a_{3+4k}, and a_{4+4k} $(0 \leq k < L/4)$. Note that the contours have rotational symmetry, but they don't meet Granlund's necessary condition for rotationally symmetric objects. This discrepancy appears frequently in the remaining plots and will be explained later.

Figs. 7-12 are for a cross-shaped contour. They correspond exactly to the plots for the squares and have many of the same properties.

There are two reasons why some of the generated shapes had rotational symmetry without meeting Granlund's condition for symmetry. The first reason is that Granlund assumed that the contour was traced only once. He also assumed that the lines of the contours did not cross over themselves as in the star.

These conditions are met when descriptors are calculated from a contour, but they may not be met when a contour is generated from an artificial set of descriptors.

A simple example of a contour being multiply traced would be a circle generated from a Fourier component other than a_1. The parametric representation of a circle consists of sine and cosine components with equal amplitudes. The component a_1 will generate a circle which will trace once around its contour in a period of 2π. The coefficient a_2 has twice as high a frequency, so it will trace around a circle twice in the period 2π. The resulting circles look alike even though they have different coefficients and, in a certain sense, different contours.

The properties of contours generated by descriptors whose only non-zero coefficients are of the form a_{i+pm} will now be examined. The symbols stand for:

i initial component, lowest frequency component
m multiples, every m^{th} component may be non-zero
p integer variable

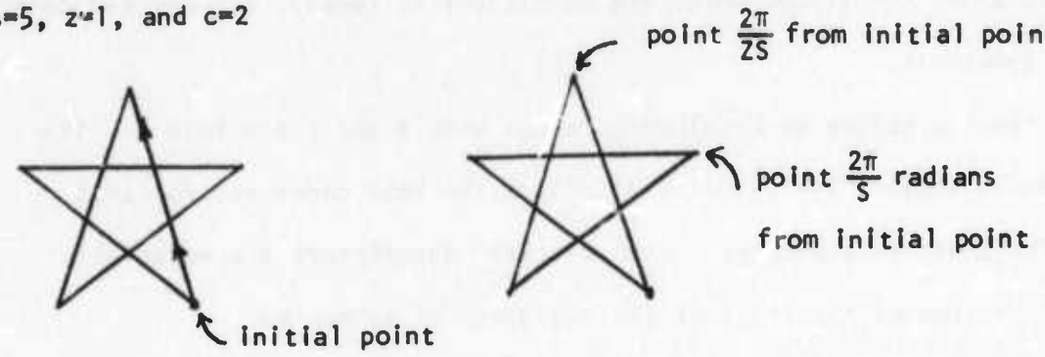
The following symbols will be useful:

s degree of rotation symmetry of the contour
z the number of times the contour is traced in a period of 2π
c the number of nonoverlapping cycles made around the center of the contour

This derivation is based on Granlund's idea of rotating the object and moving the starting point so that the effects of the two changes cancel out.

Let the initial point on a rotationally symmetric contour be moved to the next corresponding point along the contour. $2\pi/z$ is the distance to move to trace once completely around the contour, so a move of $2\pi/zs$ will result in a move to the next point which is similar to the original initial point. Note that this move is to the next similar point on the contour. This may not be

the point at an angle $2\pi/s$ radians back around the contour. For example, with s=5, z=1, and c=2



Rotating the contour by $-2\pi c/s$ radians will move the initial point back to its original location in the complex plane and will result in a shape with the same Fourier descriptors as the original.

$$a_n = a'_n = a_n e^{-j\frac{2\pi}{S}c} e^{j\frac{2\pi}{ZS}n} = a_n e^{j2\pi\left(\frac{n}{ZS} - \frac{c}{S}\right)}$$

For this equation to hold either $a_n = 0$, or $(n/zs) - (c/s) = \pm p$ for integer $p \geq 0$. Solving for n gives $n = zc \pm pzs$. This result is of the form $n = i \pm pm$ which is the type of Fourier descriptor we are interested in.

The definition of c requires that it not include the result of multiple traces (traces which write directly over previous traces). If c and s were not relatively prime then multiple tracing would result before the first c cycles were complete. Therefore c and s must be relatively prime. Therefore, the greatest common divisor of i and m is z.

Relations which must hold for rotationally symmetric contours

gcd (i, m) =	z	number of tracings
m/z =	s	degree of symmetry
i/z =	c	number of distinct cycles
i =	zc	total number of cycles

These results show how to calculate the degree of rotational symmetry, number of cycles about the center of the contour, and the number of multiple

traces around the contour, given that the contour is symmetric. Unfortunately there are no known conditions which are sufficient to require a generated contour to be symmetric.

These results reduce to Granlund's result when z and c are both 1. The equations above explain the symmetry of all of the test cases run for this project. The contours generated by the circle's descriptors are especially interesting because of the range of their degrees of symmetry.

REFERENCES

[1] G. H. Granlund, "Fourier Preprocessing for Hand Print Character Recognition," IEEE Trans. on Comp., Vol. c-21, Number 2, February 1972.

[2] C. T. Zahn and R. Z. Roskies, "Fourier Descriptors for Plane Closed Curves," IEEE Trans. on Comp. Vol. c-21, Number 3, March 1972.

[3] E. Persoon and K. S. Fu, "Sequential Decision Procedures with Prespecified Error Probabilities and Their Applications," Purdue University, TR-EE 74-30, August 1974.

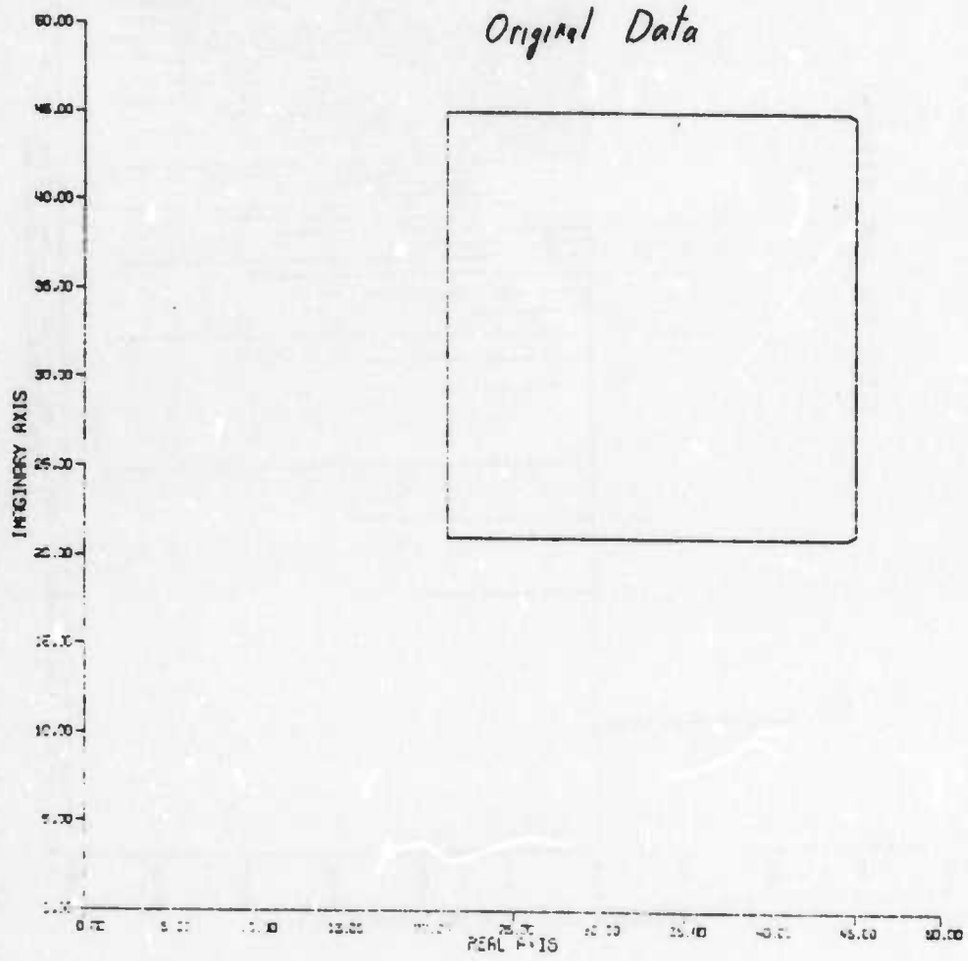


Fig. 1 Contour of the input data

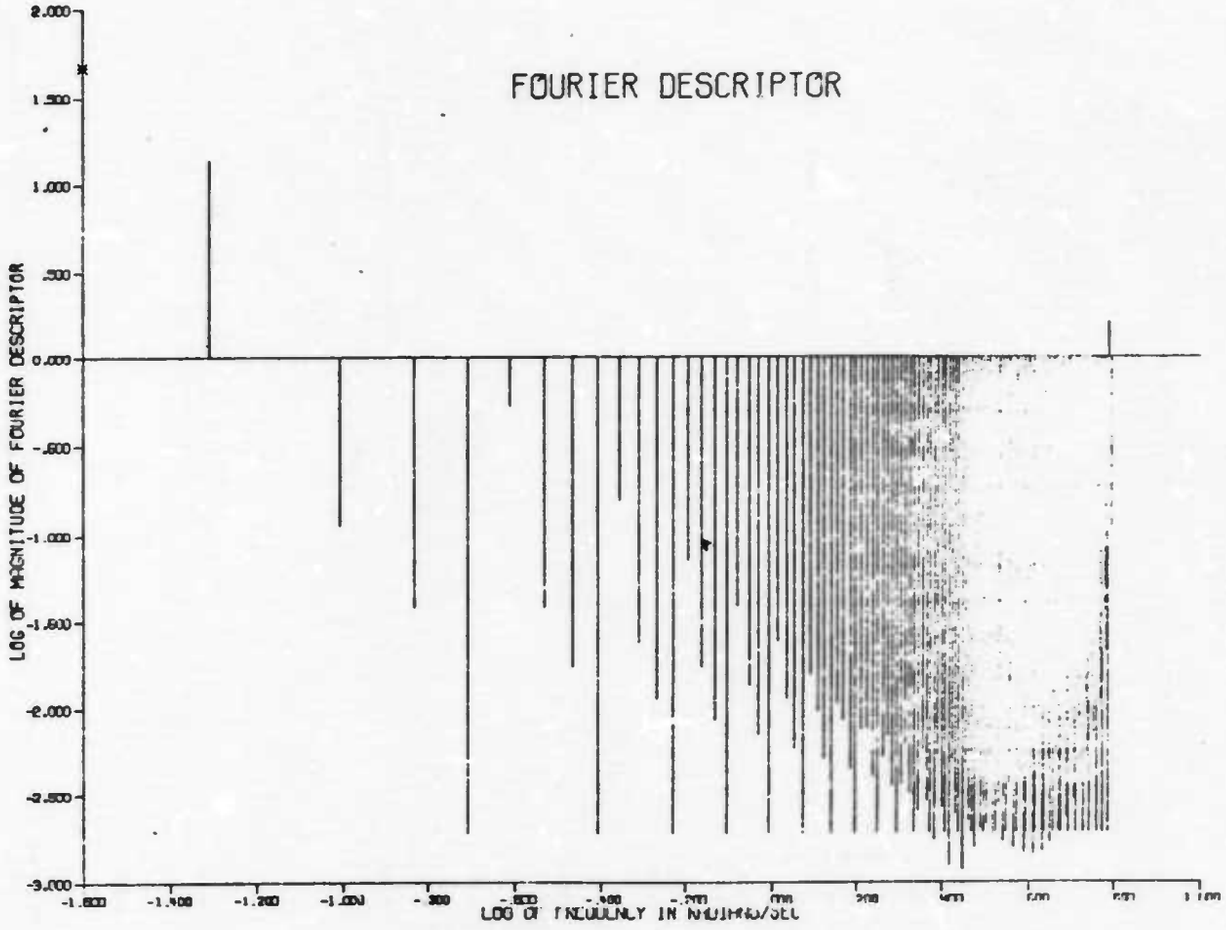


Fig. 2a Magnitudes of the Fourier Descriptor coefficients (log-log scales). The * on the Y axis represents the magnitude of a_o.

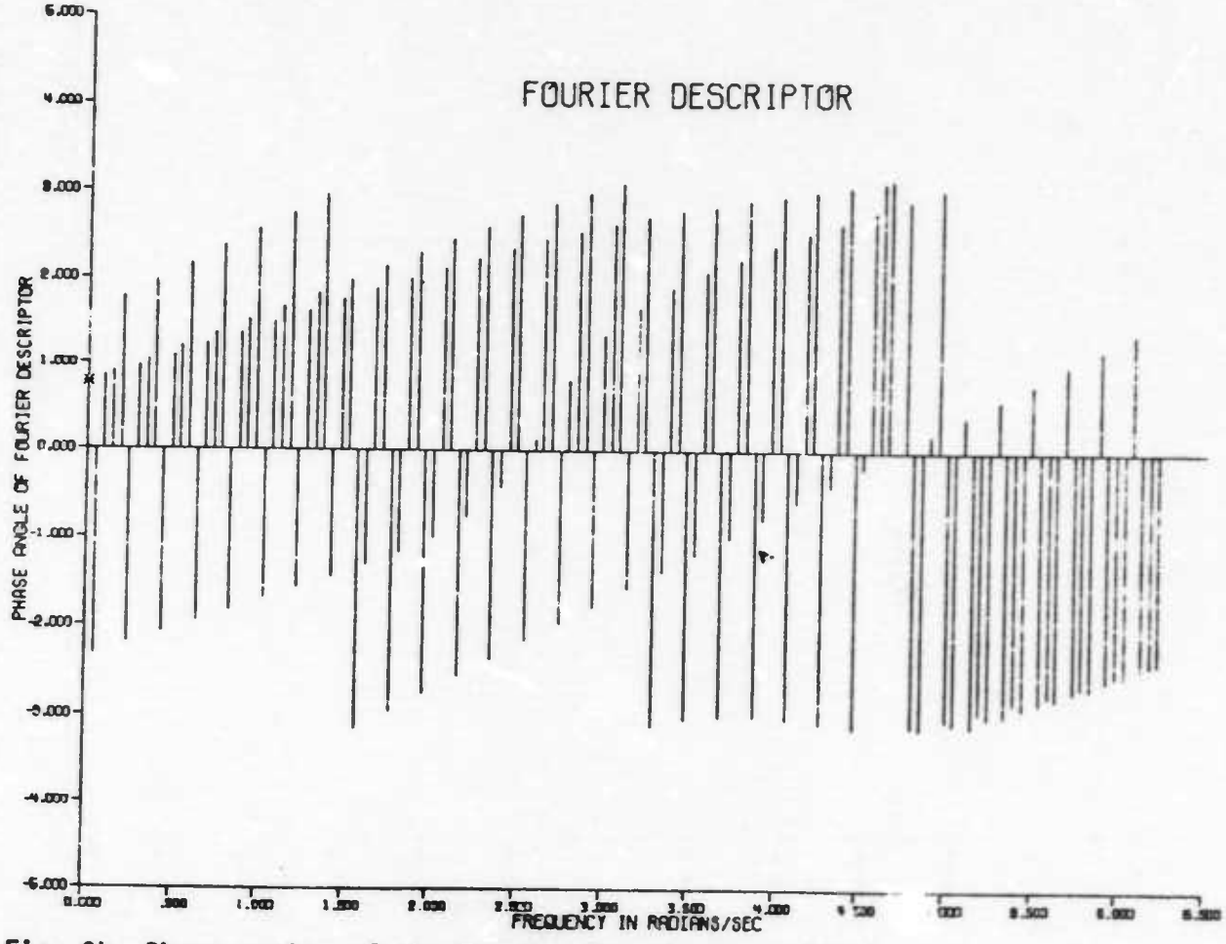


Fig. 2b Phase angles of the Fourier Descriptor coefficients.

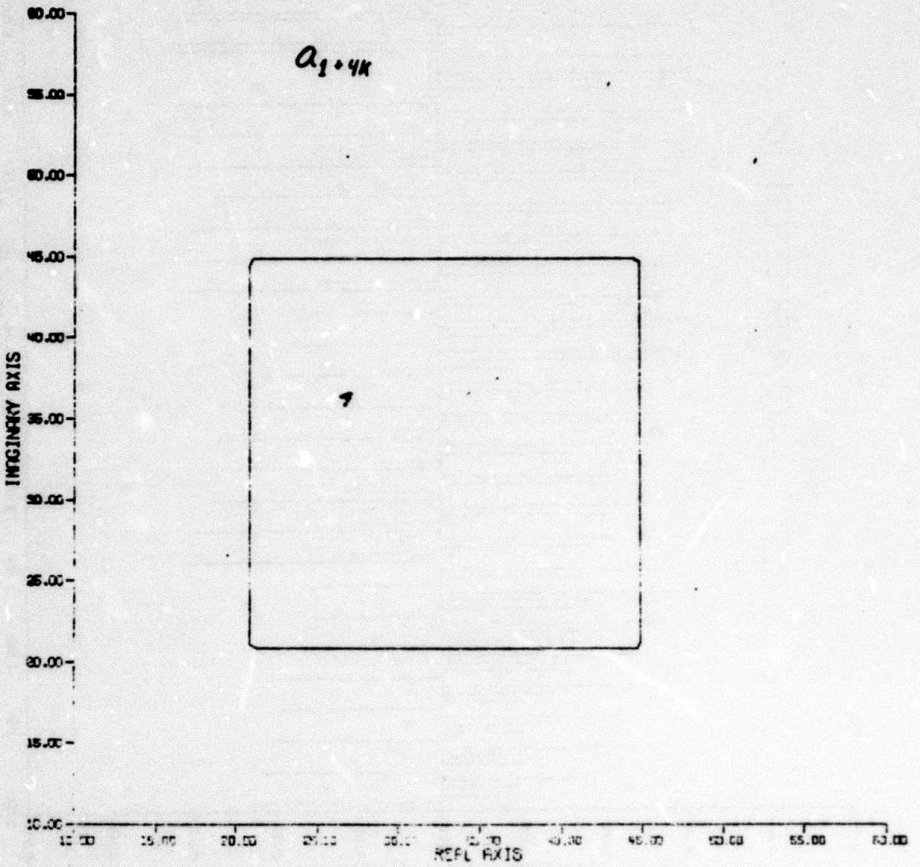


Fig. 3 The contour generated by a_o and coefficients of the form a_{1+4K} where K is an integer. The other coefficients were set equal to zero.

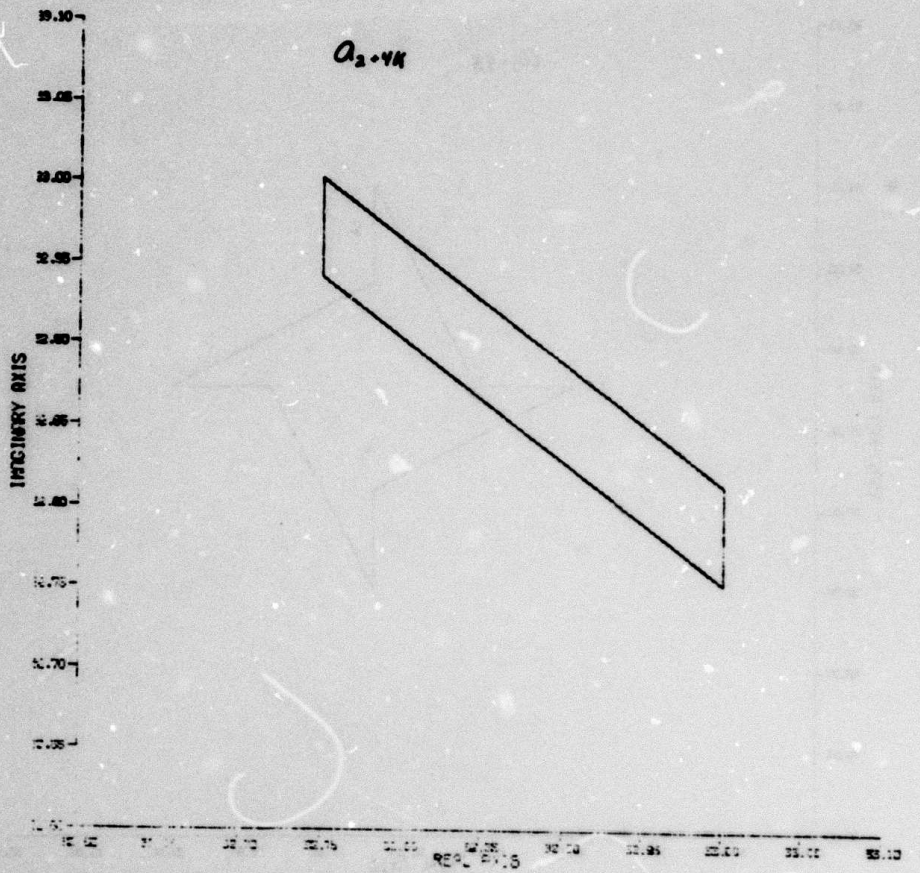


Fig. 4 The contour generated by a_o and coefficients of the form a_{2+4K} where K is an integer. The other ocefficients were set equal to zero.

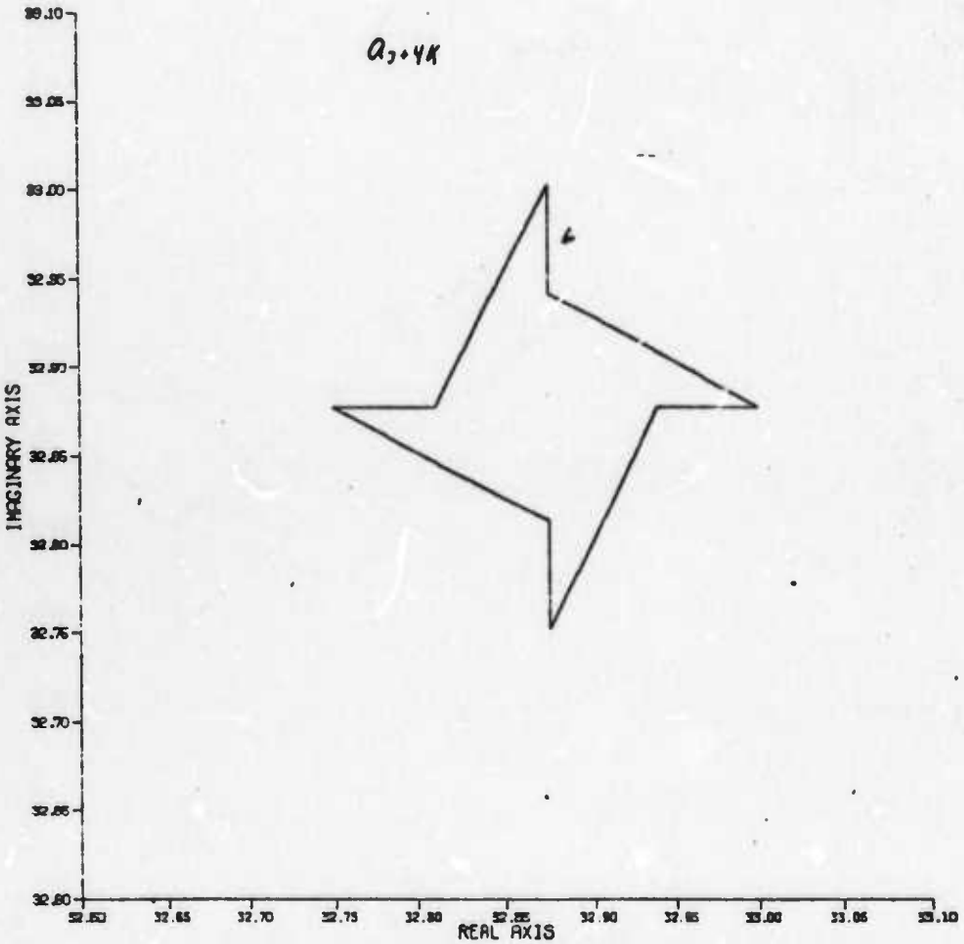


Fig. 5 The contour generated by a_o and coefficients of the form a_{3+4K} where K is an integer. The other coefficients were set equal to zero.

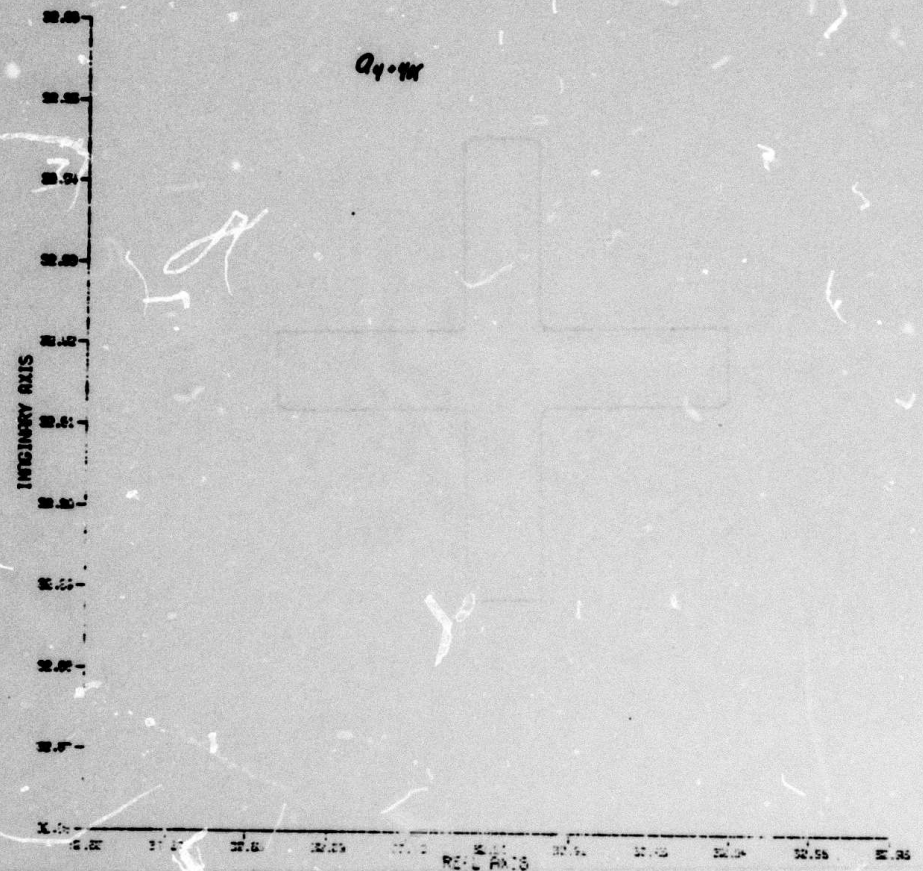


Fig. 6 The contour generated by a_o and coefficients of the form a_{4+4K} where K is an integer. The other coefficients were set equal to zero.

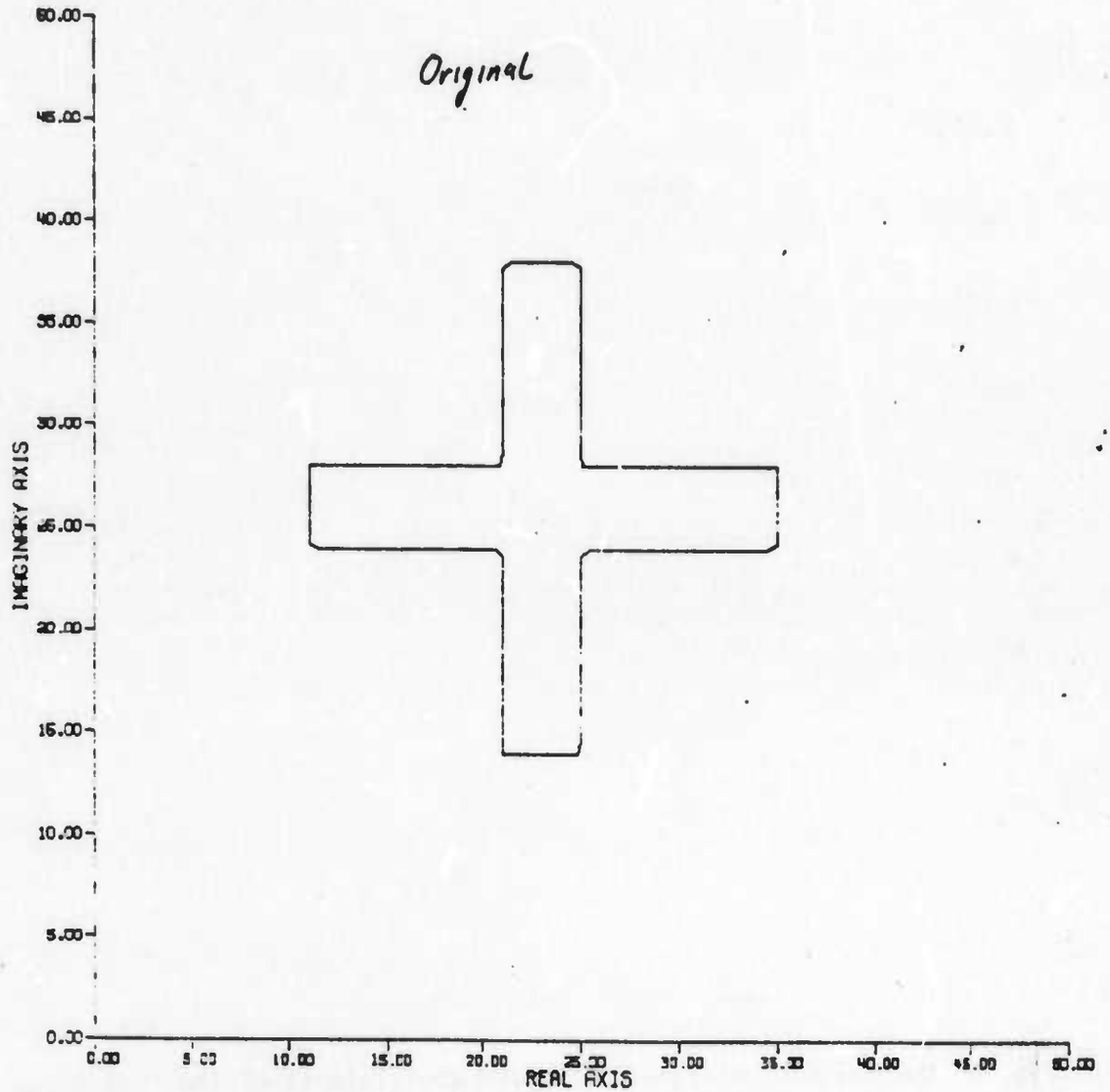


Fig. 7 Contour of the input data

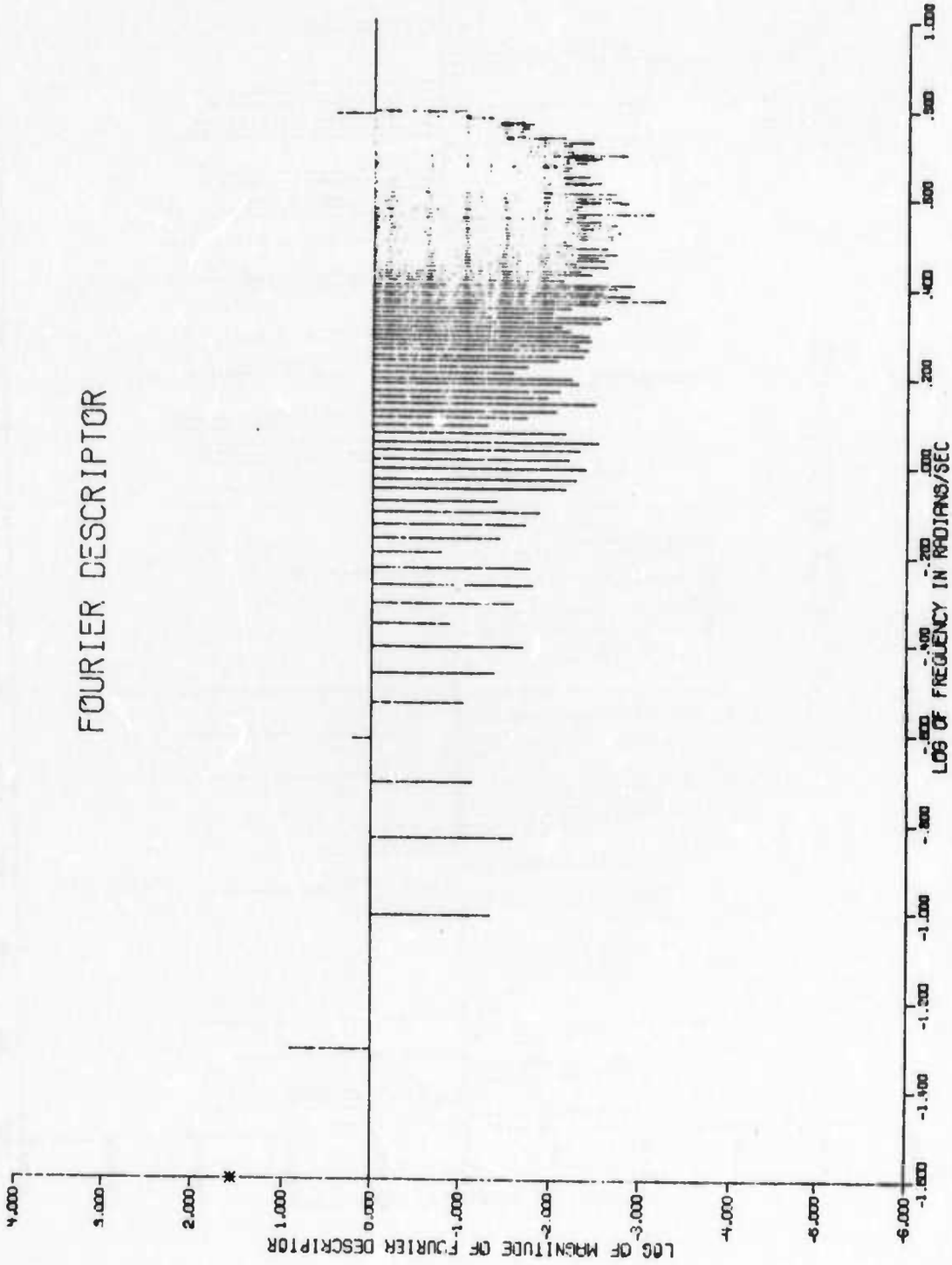


Fig. 8a Magnitudes of the Fourier Descriptor coefficients (log-log scales). The * on the Y axis represents the magnitude of a_o.

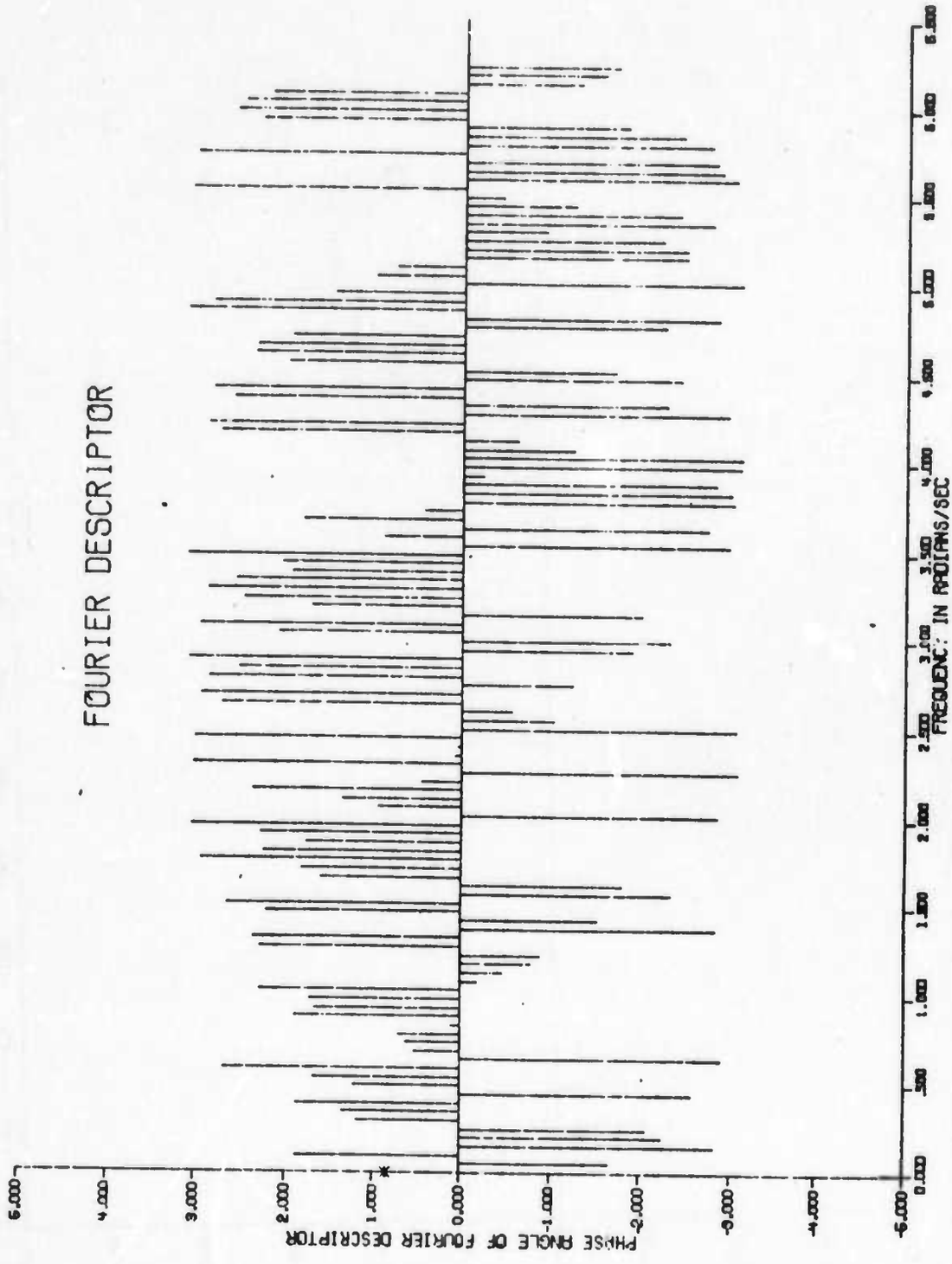


Fig. 8b Phase angles of the Fourier Descriptor coefficients.

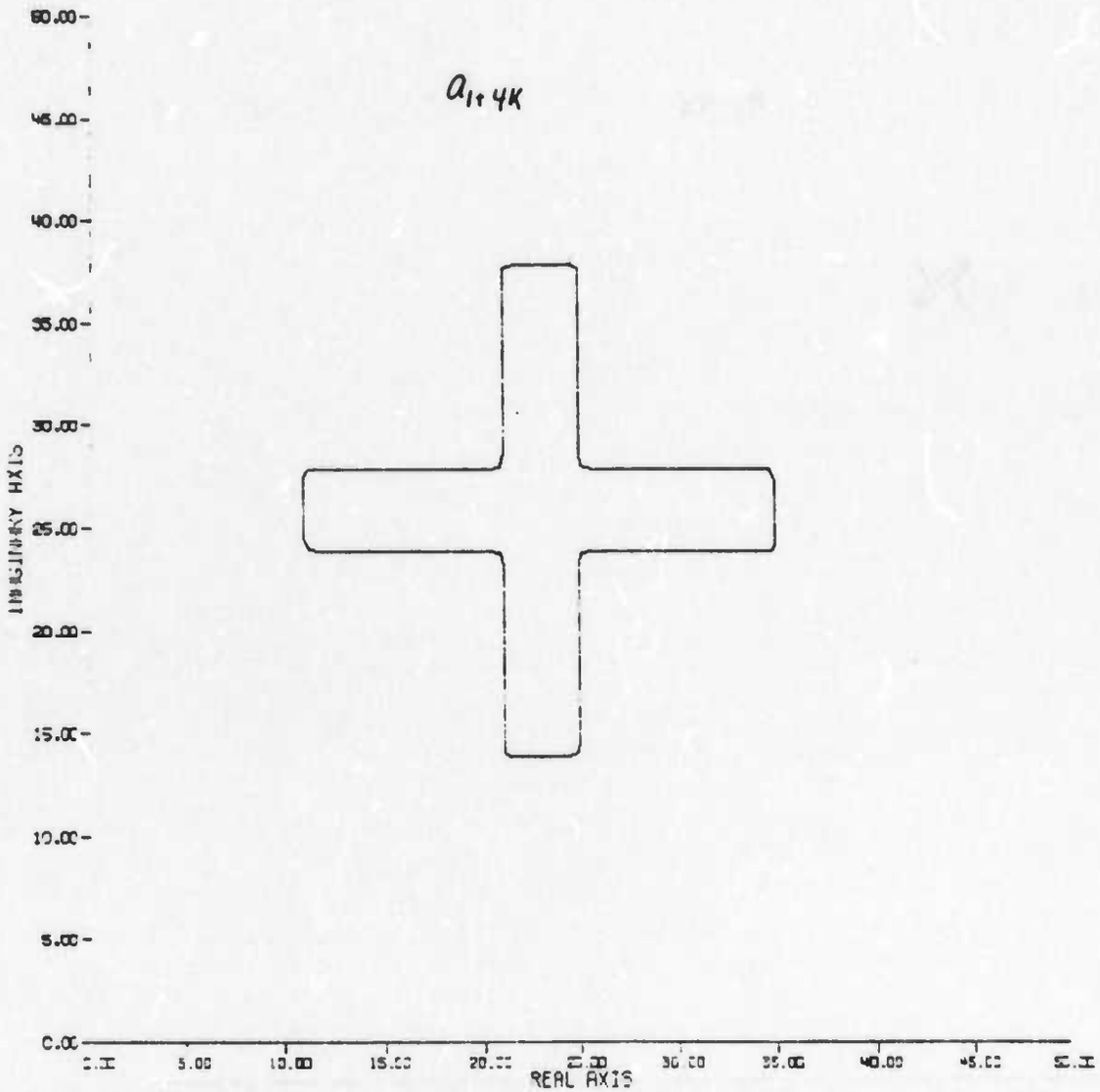


Fig. 9 The contour generated by a_o and coefficients of the form a_{1+4K} where K is an integer. The other coefficients were set equal to zero.

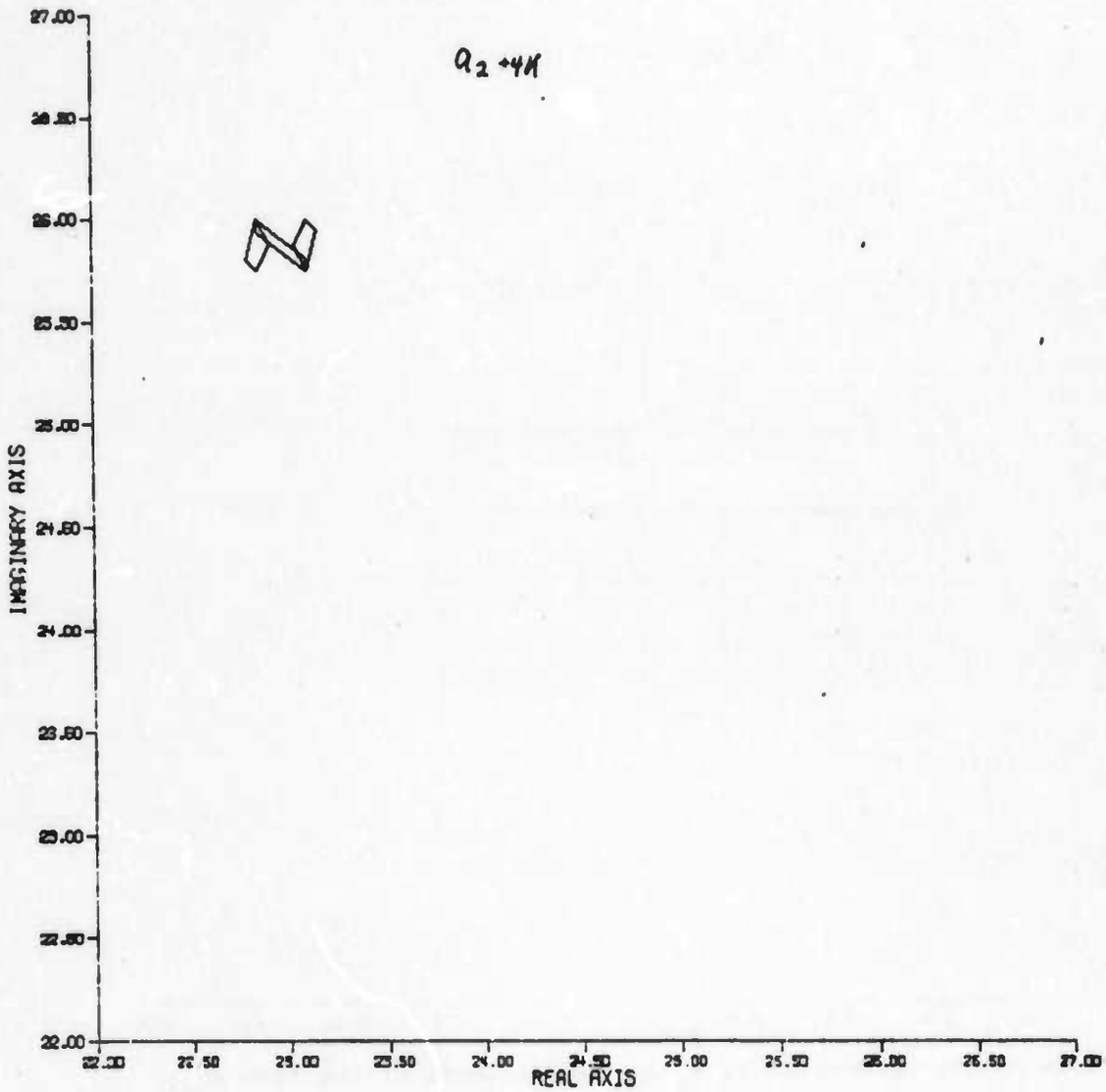


Fig. 10 The contour generated by a_o and coefficients of the form a_{2+4K} where K is an integer. The other coefficients were set equal to zero.

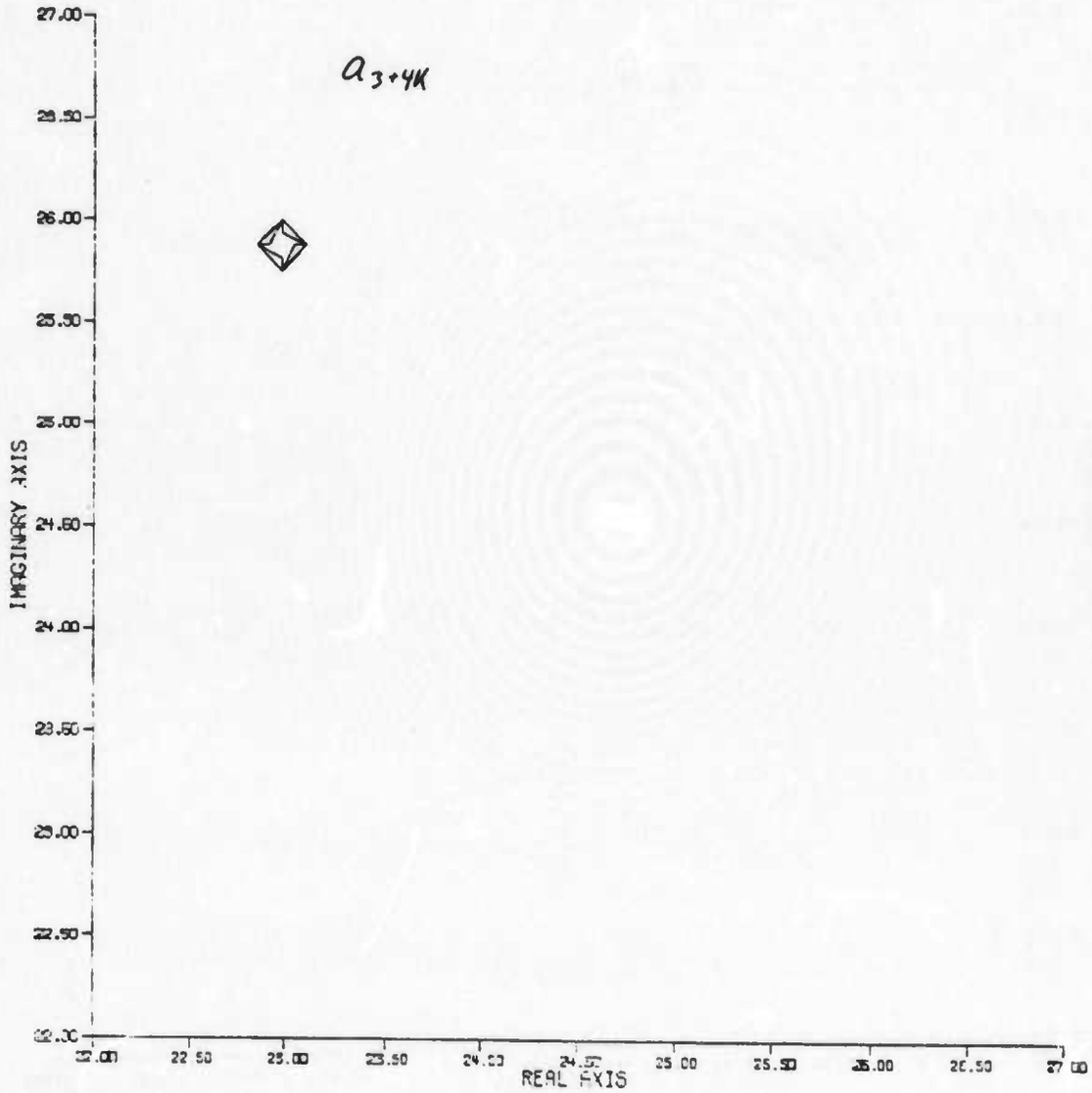


Fig. 11 The contour generated by a_o and coefficients of the form a_{3+4K} where K is an integer. The other coefficients were set equal to zero.

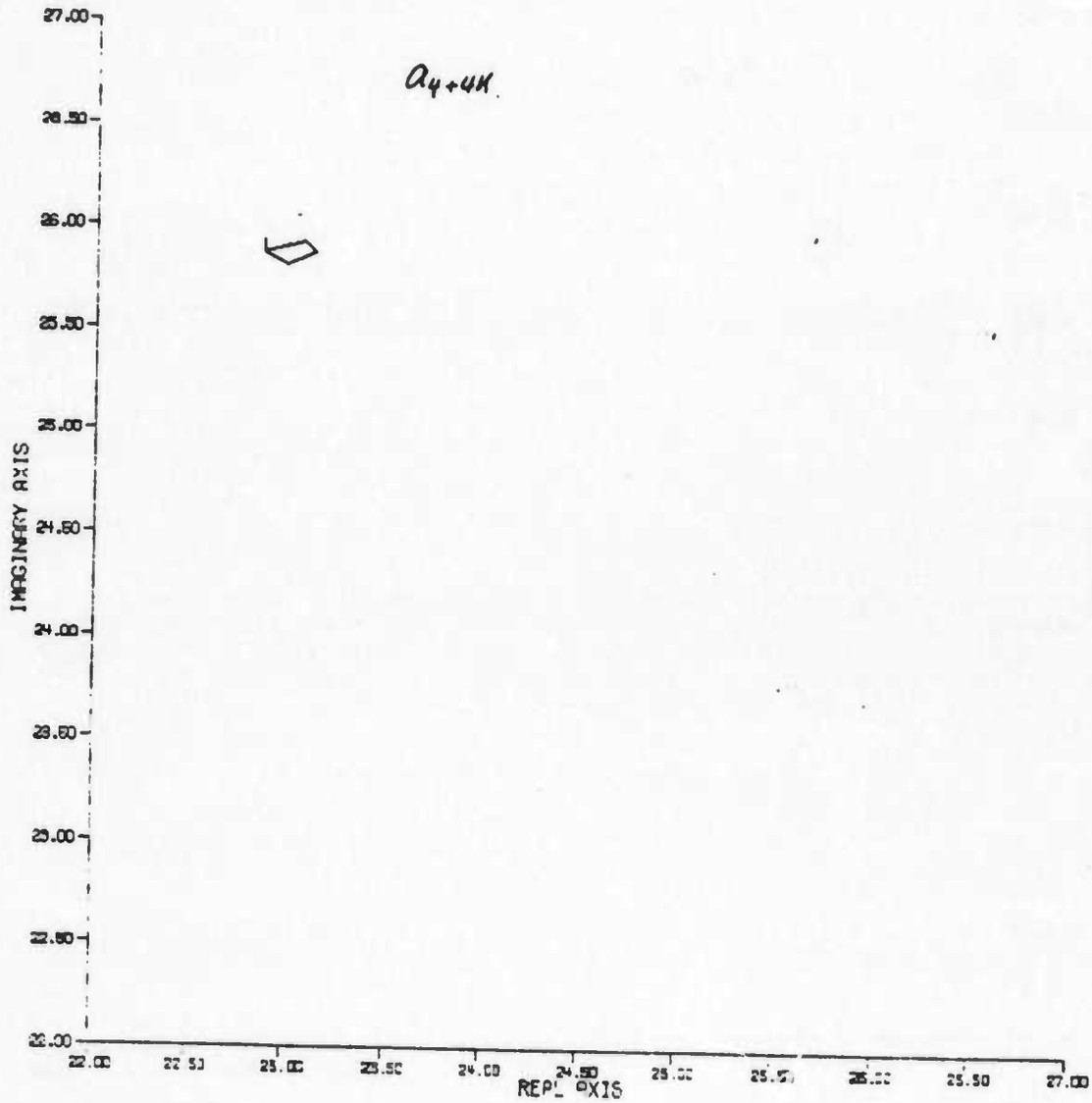


Fig. 12 The contour generated by a_o and coefficients of the form a_{4+4K} where K is an integer. The other coefficients were set equal to zero.

SELECTION OF SPECTRAL BANDS FOR USE IN MULTISPECTRAL SCENE ANALYSIS

E. Wiswell and G. Cooper

I. INTRODUCTION

Selection of spectral bands in multispectral scanners has to the present time been based largely on available energy considerations. Hence, selection of spectral bands with consideration to numerical scene analysis has to date been largely empirical. Therefore, this investigation is directed toward selection of bands in multispectral scanners quantitatively with regard to the use of the data for numerical processing. To optimally select spectral bands, a criteria of optimality must accordingly be chosen and evaluated. Thus, this project is concerned with optimal multispectral band selection with regards to numerical scene analysis and a suitable optimality criteria.

II. AREA OF INVESTIGATION

Presently, this investigation is attempting to properly apply information theoretic concepts to the problem of multispectral band selection.

A first step is to consider the spectral response as a random process in wavelength. Hence, represent the spectral response as:

$$y(\lambda) = s(\lambda) + n(\lambda)$$

Here $s(\lambda)$ is the true spectral response. $n(\lambda)$ is a term that includes noise (i.e. thermal noise) as well as any interference. An example of interference might be bare soil between rows of crop vegetation early in the growing season. It is important to carefully consider the consequences of declaring part of the received spectral response as noise when it may actually contain information useful for numerical scene analysis. Perhaps it would be useful

to decompose $n(\lambda)$ into a noise term and an interference term. This thought remains to be investigated.

A spectral interval (λ_1, λ_2) where λ is wavelength is considered. The interval is to be subdivided into several spectral bands. It is proposed to compute the mutual information between spectral bands and to choose a subset of spectral bands which maximizes the mutual information between the remaining bands.

In order to compute the mutual information between spectral bands, the autocorrelation functions of the signal $R_s(\lambda_1, \lambda_2)$ and $R_n(\lambda_1, \lambda_2)$ (corresponding to the $n(\lambda)$ term) are needed. Therefore, an effort is currently under way to estimate the autocorrelation functions from representative multispectral data obtained at Purdue's Laboratory for Applications of Remote Sensing (LARS).

A problem related to the computation of mutual information is consideration of the effects of width of the spectral bands. Another related problem concerns the relationship of the spectral bands to temporal and spatial information. The effects of a set of multispectral bands and their width will ultimately be related to spatial resolution for example. And, of course, there is no reason to suppose that spectral information will be constant temporally.

III. FUTURE INVESTIGATION

There are two main topics to be pursued. First, the information theoretic approach will be continued. Secondly, it may be useful to pursue a decision theoretic approach. The motivation for this that one may be able to choose a cost function that reflects the application for which the data is to be used (i.e., classification). The extent to which any decision theoretic approach differs from previous techniques (i.e., pattern recognition approaches) must be determined.

IV. COMMENTS

Due to the fact that this investigation is in its initial stages, application oriented results have yet to be achieved. However, a theoretical approach that is perceived as viable has been formulated. The experimental demonstration of this approach is to be pursued.

ITERATIVE IMAGE RESTORATION, III

T. S. Huang, S. Berger, and M. Kaveh

I. INTRODUCTION

The restoration of images degraded by linear space-variant degradations can be accomplished through the use of the projection algorithm. This algorithm has been applied to both one and two-dimensional signals in order to examine its capabilities. The current results indicate that the method may be quite valuable in the restoration of images.

II. INVESTIGATION OF THE ALGORITHM

The projection algorithm is basically an iterative method for the solution of a system of equations

$$g_1 = a_{11} f_1 + a_{12} f_2 + \ldots + a_{1n} f_n$$

$$g_2 = a_{21} f_1 + \ldots \qquad + a_{2n} f_n$$

$$\vdots$$

$$g_m = a_{m1} f_1 + \ldots \qquad + a_{mn} f_n$$

where n may not be equal to m. The solution is obtained by successive projections onto planes in hyperspace. For the case where a unique solution does exist, the algorithm will converge to the point of intersection of the hyperplanes. If the planes do not intersect at a single point, the algorithm will converge to a point which may be a useful approximation in a restoration sense.

In order to apply the algorithm, one must consider the degradation process in terms of discrete components. A one-dimensional signal is divided

into evenly spaced samples, and a two-dimensional image is treated as an array of picture elements. The degradation process consists of a known linear degradation and additive Gaussian noise.

The performance of the algorithm was investigated for several cases. A one-dimensional signal was degraded by a low-pass filter. The original signal f(x) consists of two unit pulses of five points each, separated by two zero points (Fig. 1). The degradation is effected in the Fourier transform domain. The discrete Fourier transform (DFT) of f(x) is obtained by a fast Fourier transform routine. A triangular low-pass filtering operation (with a cutoff frequency equal to .22 times the maximum frequency) yields the degraded signal $g_l(x)$ (Fig. 2). The result after the addition of Gaussian noise is denoted by g(x) (Fig. 3). The initial guess for the algorithm is taken to be g(x). The algorithm was tested for different restrictions on the successive iterations $f^l(x)$.

The projection algorithm was compared with the least-squares inverse filter, whose frequency response is given by

$$Q(u) = \frac{H^*(u) \cdot S_f(u)}{|H(u)|^2 \cdot S_f(u) + S_n(u)}$$

where H(u) is the degradation filter, $S_n(u)$ and $S_f(u)$ are the spectral densities of the noise and signal, respectively.

A two-dimensional signal was also used to test the projection algorithm. An image was degraded by an averaging operation. Each picture element was represented by an integer from 0 to 255, which is 8 bit quantization. The 128x128 image arrays were stored on magnetic tape in a compacted format. The Gould electrostatic plotter was used to obtain half-tone reproduction of the images.

III. RESULTS

The restoration was implemented on a CDC 6500 computer. For the one-dimensional case, the additive Gaussian noise had a standard deviation of $\sigma = 0.05$, which yields a signal-to-noise ratio of $20 \log_{10} \frac{1}{.05} = 26$ dB.

The result of the least squares inverse filter is shown in Fig. 4. The spectral density of the signal $S_f(u)$ was assumed to be a raised cosine. The filter is not able to resolve the original two peaks.

The result of the projection algorithm, when no restriction is placed on the signal, is given in Fig. 5a and 5b for 10 and 40 iterations, respectively. The peaks are resolved; however, there is considerable error outside the interval containing the peaks. When the signal after each iteration is restricted to only positive values, $f^i(x) \geq 0$, the resolution of the two peaks is greatly improved. And the overall error is reduced considerably. The results after 10 and 40 iterations is given in Fig. 6a,b.

The projection algorithm is versatile in that it is capable of utilizing a priori information about the signal. If the values outside the region of the two rectangular pulses are known to be zero, the algorithm can use this information. The result of setting $f(x) = 0$ after each subiteration for points outside the interval is shown in Fig. 7a, b for 1 and 40 iterations. The first iteration is shown because the algorithm was able to produce two peaks, and this was not the case before. The fourtieth iteration is actually not significantly different than the tenth iteration, which is not shown. A more proper way to utilize a priori information would be to incorporate it into the structure of the algorithm. This approach will be considered in the near future.

The computation time required for the inverse filter is about 0.8 seconds. One iteration of the projection algorithm takes approximately 0.4 seconds.

The results for the two-dimensional images indicate that the projection method can improve on the degraded image, but the effects of noise dominate the image after a certain number of iterations. The optimum number of iterations is a subjective quantity. The computation time required for one iteration on a 128x128 image array is about 12 seconds.

IV. CONCLUSIONS

The results show that the projection algorithm can be a useful image restoration technique. The method is capable of incorporating restrictions on the signal and a priori knowledge. The method has been applied successfully to one and two-dimensional signals.

Future work in this area will involve the application of the algorithm to images with more involved types of degradation. The effect of error in the coefficients of the degrading transformation will be investigated.

REFERENCES

[1] K. Tanabe, "Projection Method for Solving a Singular System of Linear Equations and its Applications," Numer. Math., Vol. 17, pp. 203-214, 1971.

[2] T. S. Huang and P. Barker, "Iterative Image Restoration," Summary Report of Research, TR-EE 74-22, Purdue University, School of Electrical Engineering, West Lafayette, Indiana, June 1974.

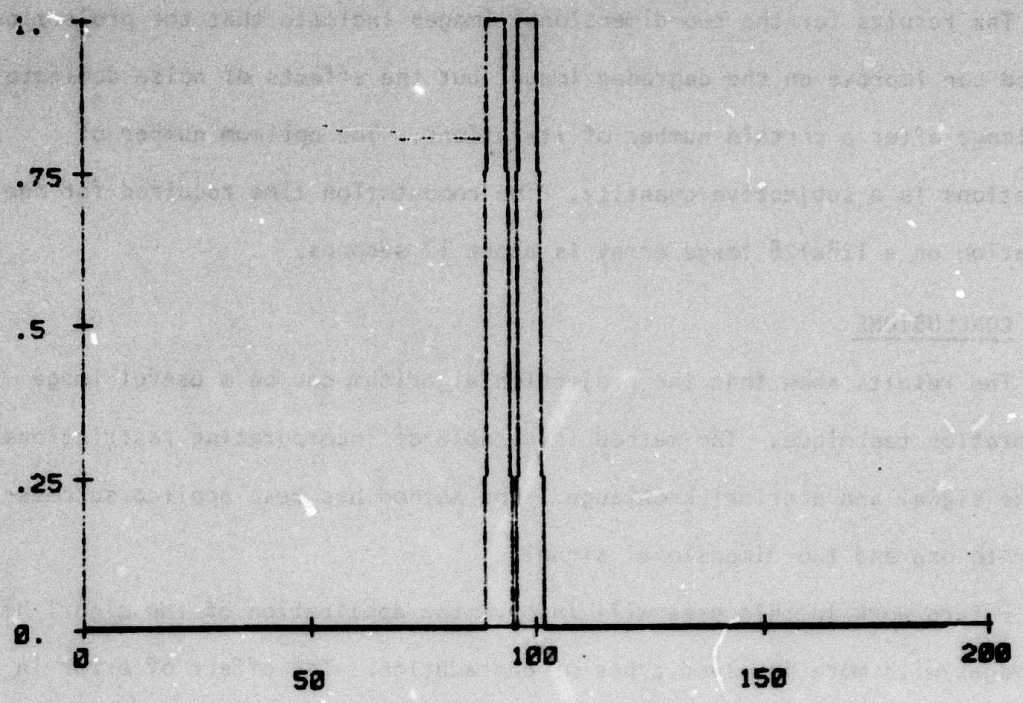


Fig. 1 Original signal

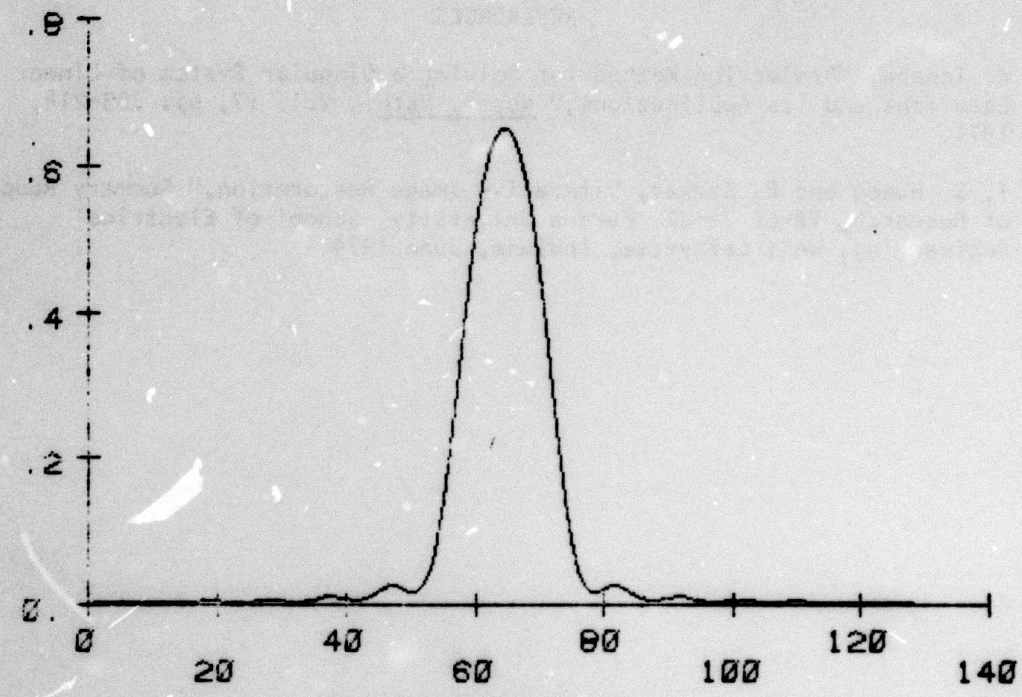


Fig. 2 Smeared signal

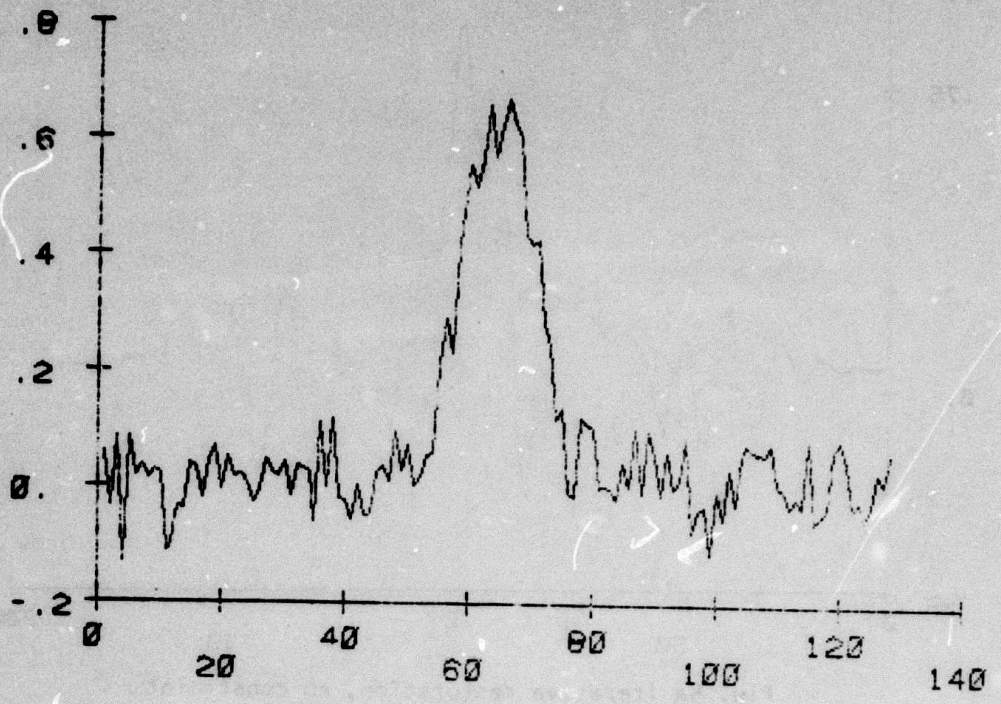


Fig. 3 Noisy smeared signal

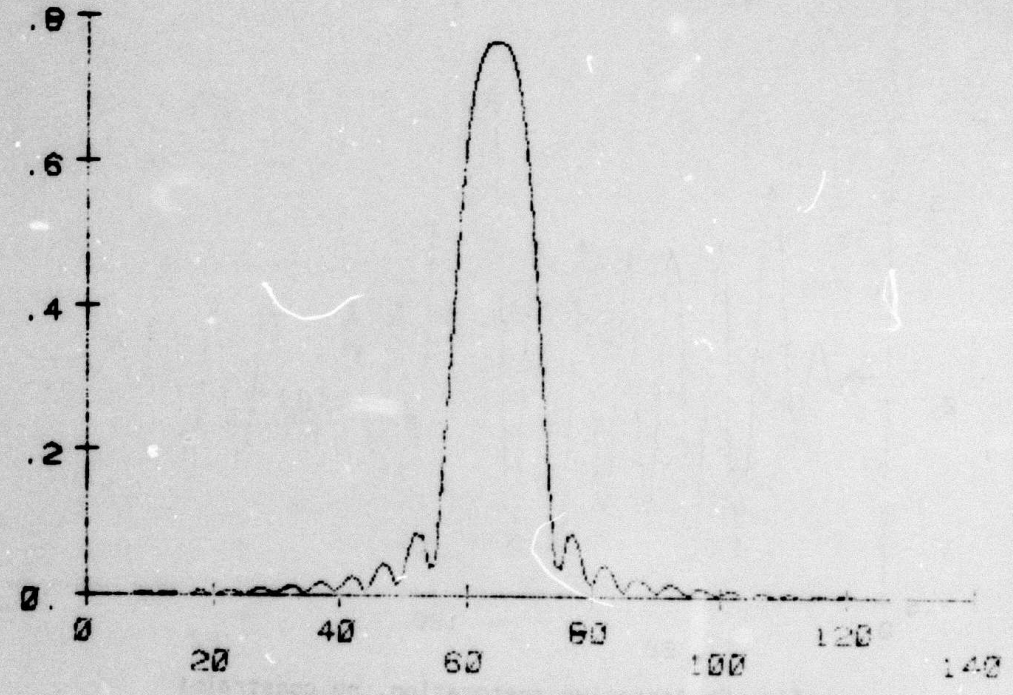


Fig. 4 Restoration by least-square inverse filter

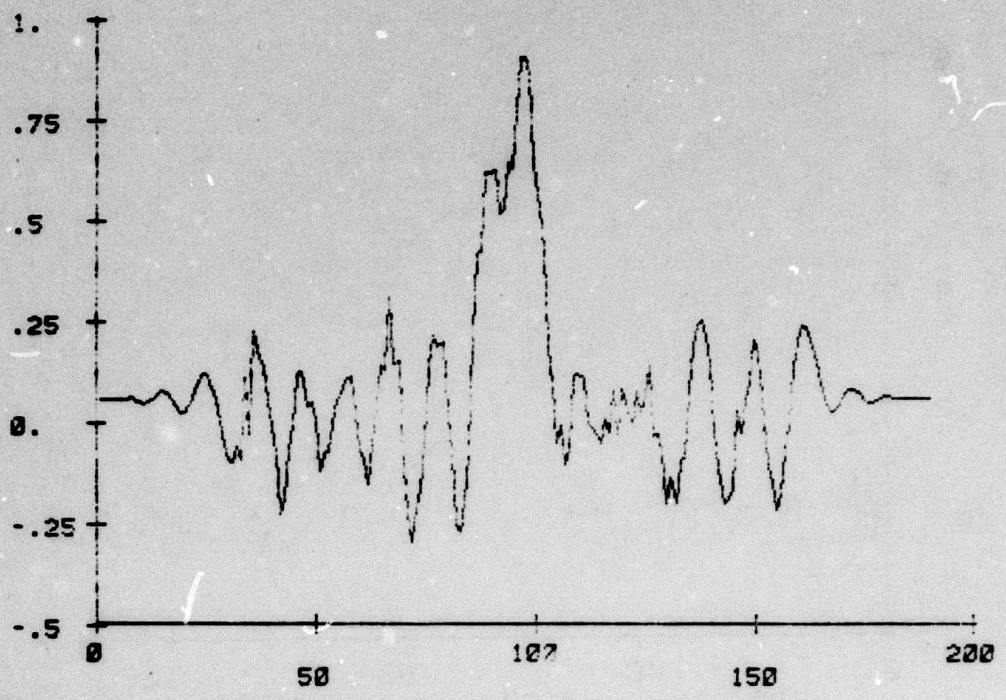


Fig. 5a Iterative restoration, no constraint
(10 iterations)

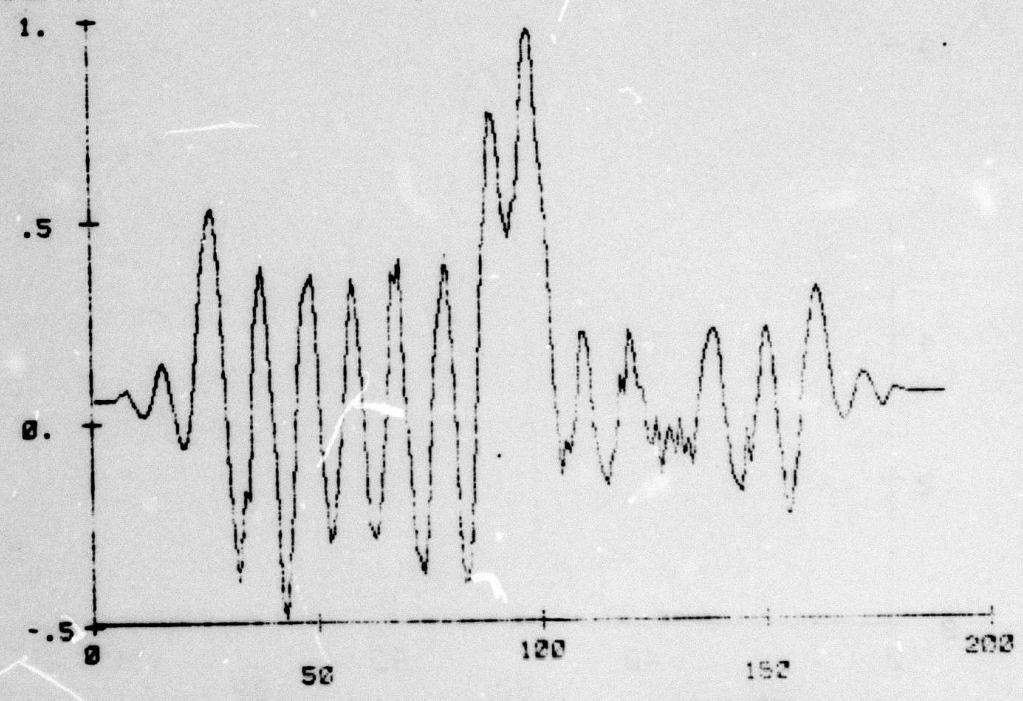


Fig. 5b Iterative restoration, no constraint
(40 iterations)

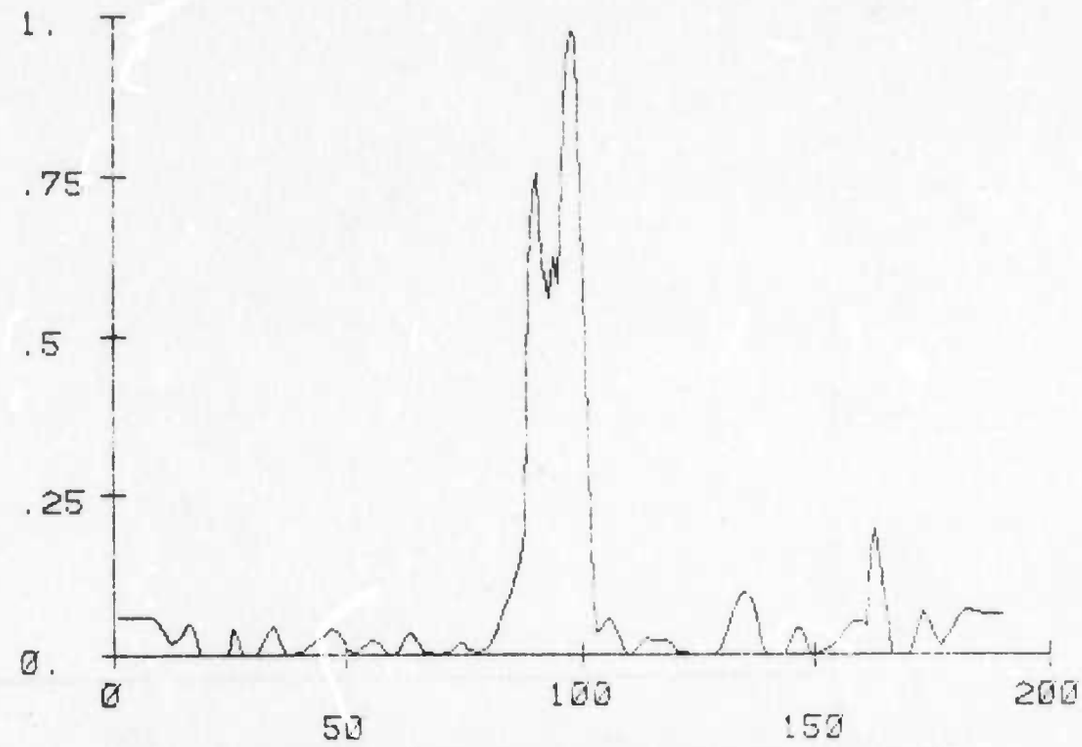


Fig. 6a Iterative restoration, positivity constraint
(10 iterations)

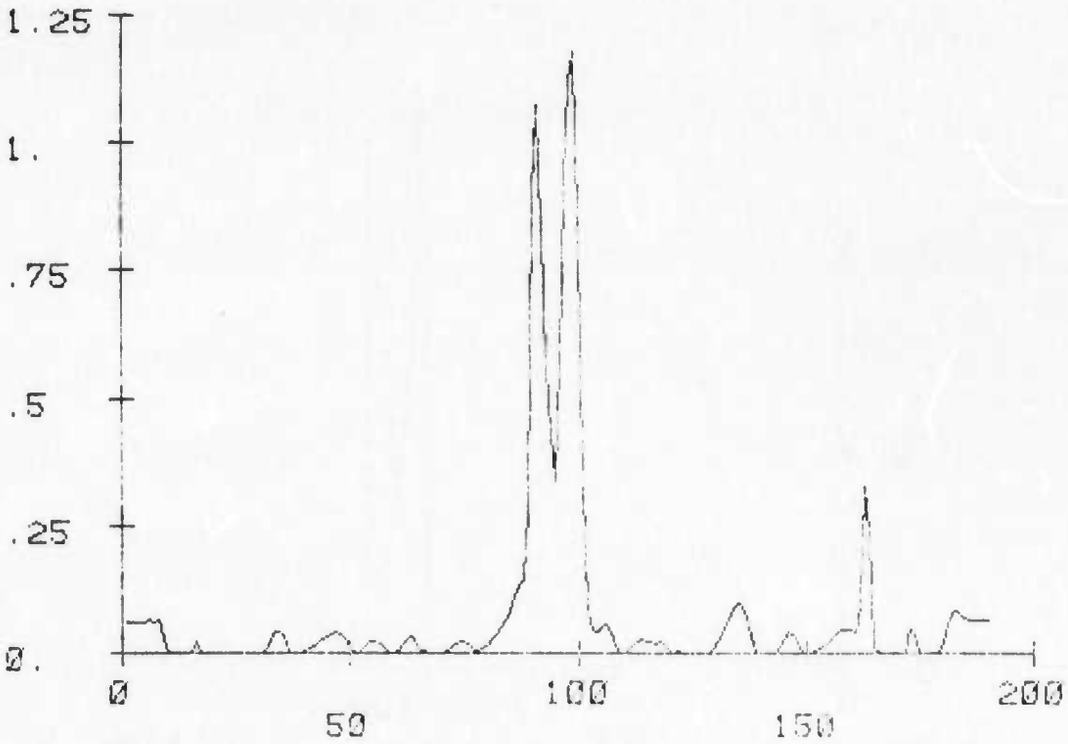
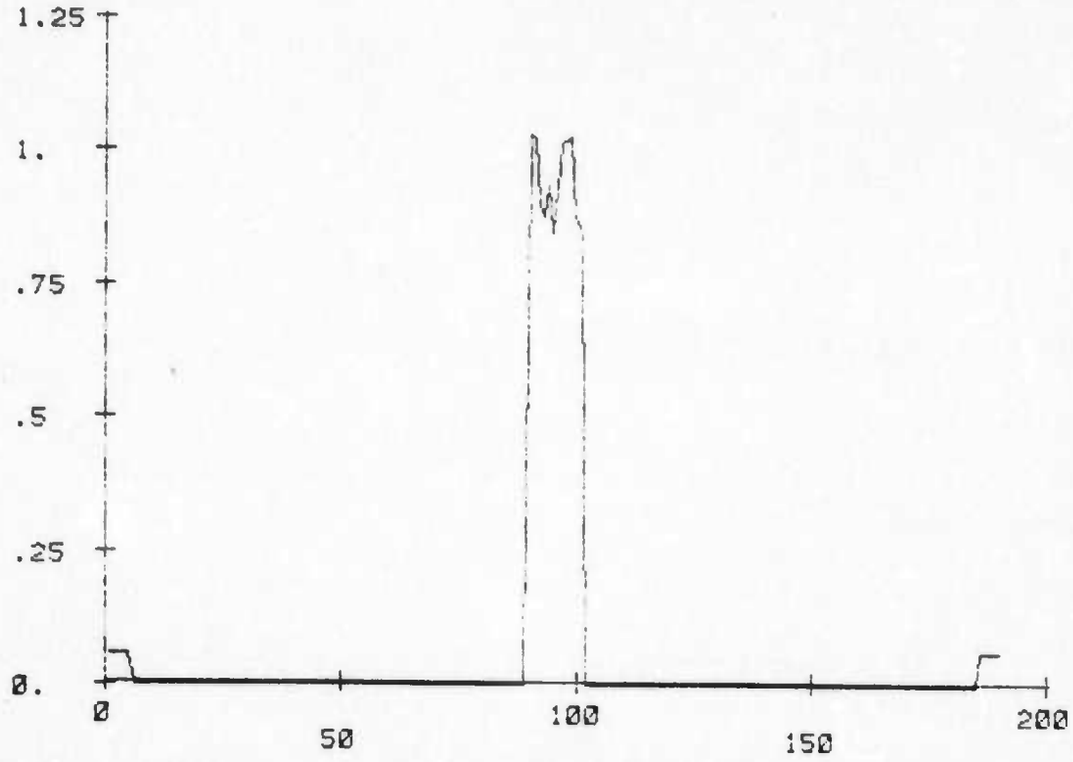


Fig. 6b Iterative restoration, positivity constraint
(40 iterations)



Fit. 7a Iterative restoration, shape constraint
(1 iteration)

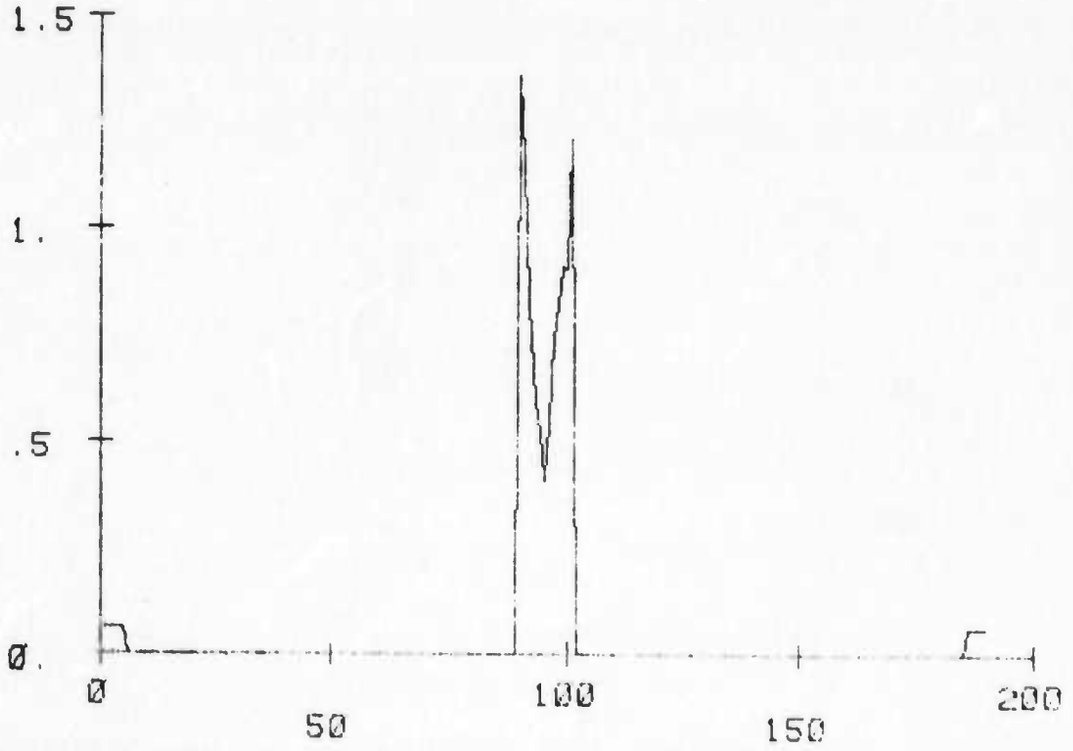


Fig. 7b Iterative restoration, shape constraint
(40 iterations)

SATELLITE IMAGE NOISE REMOVAL

O. R. Mitchell and P. L. Chen

INTRODUCTION

We are using homomorphic filtering techniques to remove multiplicative noise effects such as cloud and atmospheric disturbances in ERTS imagery. A simulated implementation using simple computer generated pictures has been performed. The picture shown in Figure 1 was multiplied by various types of noise and then filtered to recover the original picture.

RANDOM AND LOW-FREQUENCY NOISE

Low frequency noise was generated by passing computer generated random noise through a two dimensional low-pass filter. The original noise was uncorrelated (white) and uniformly distributed between 0 and 1. The picture size is 64x64 pixels. An ideal rectangular low-pass filter was used which passed the first 8 frequency components in both the horizontal and vertical directions and rejected the upper 24 frequency components in each direction.

Figures 2(a) and 2(b) show the histograms of the random noise and the low-pass noise, respectively. The intensity picture and the autocorrelation of the low-pass noise are shown in Figures 2(c) and 2(d), respectively.

HOMOMORPHIC FILTERING PROCESS

The two dimensional filtering to reduce the noise effect is accomplished by weighting each point in the Discrete Fourier Transform of the logarithm of the picture intensities by a noncausal filter function of the form $\frac{S_i}{S_i+N_i}$ where S_i and N_i are the values of the signal and noise power spectrums at the particular frequency point being considered. Following the filtering process we take the inverse DFT and then exponentiate to obtain the filtered picture.

Fig. 1 Computer Generated Picture (64x64)

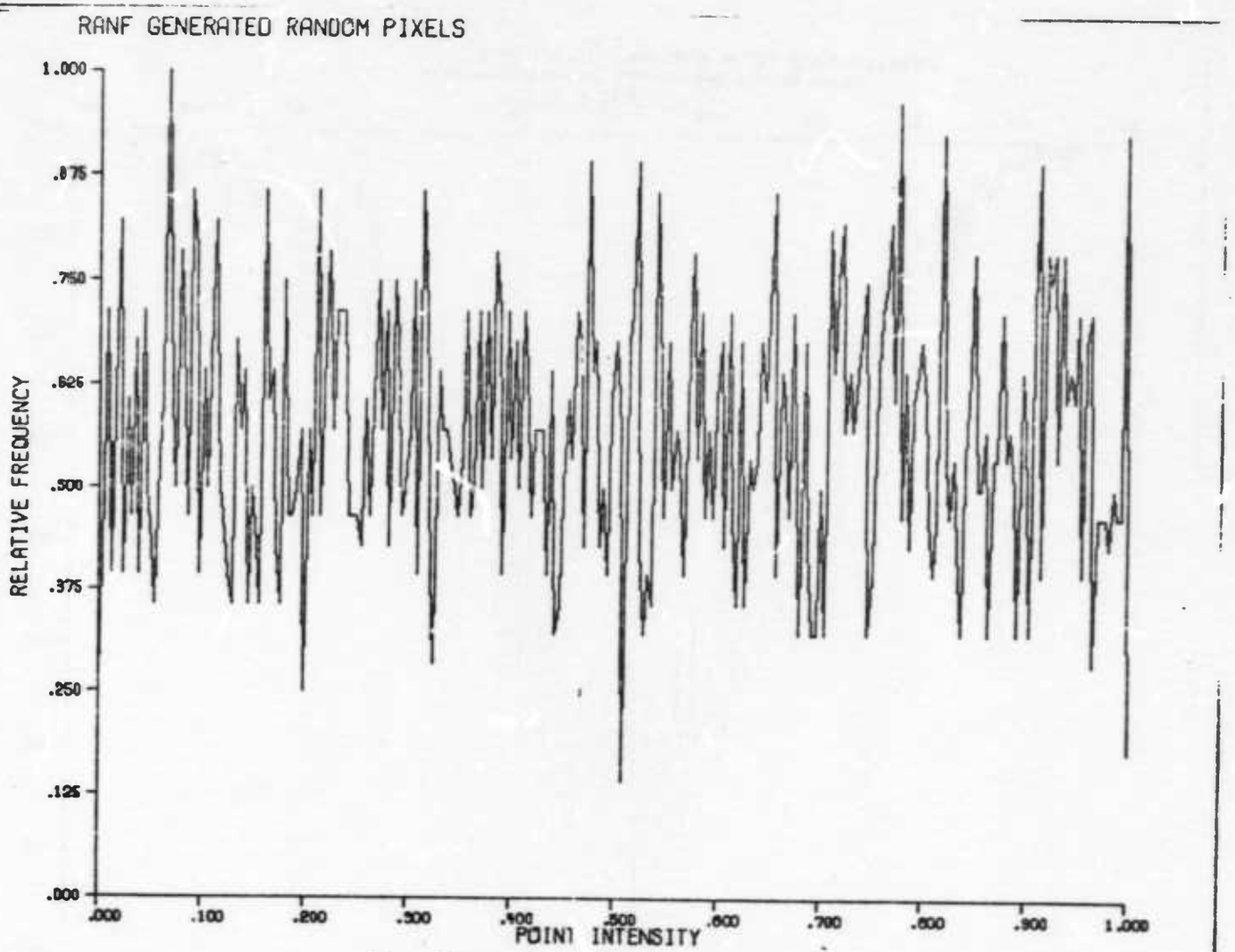


Fig. 2(a) Histogram of the Random Noise

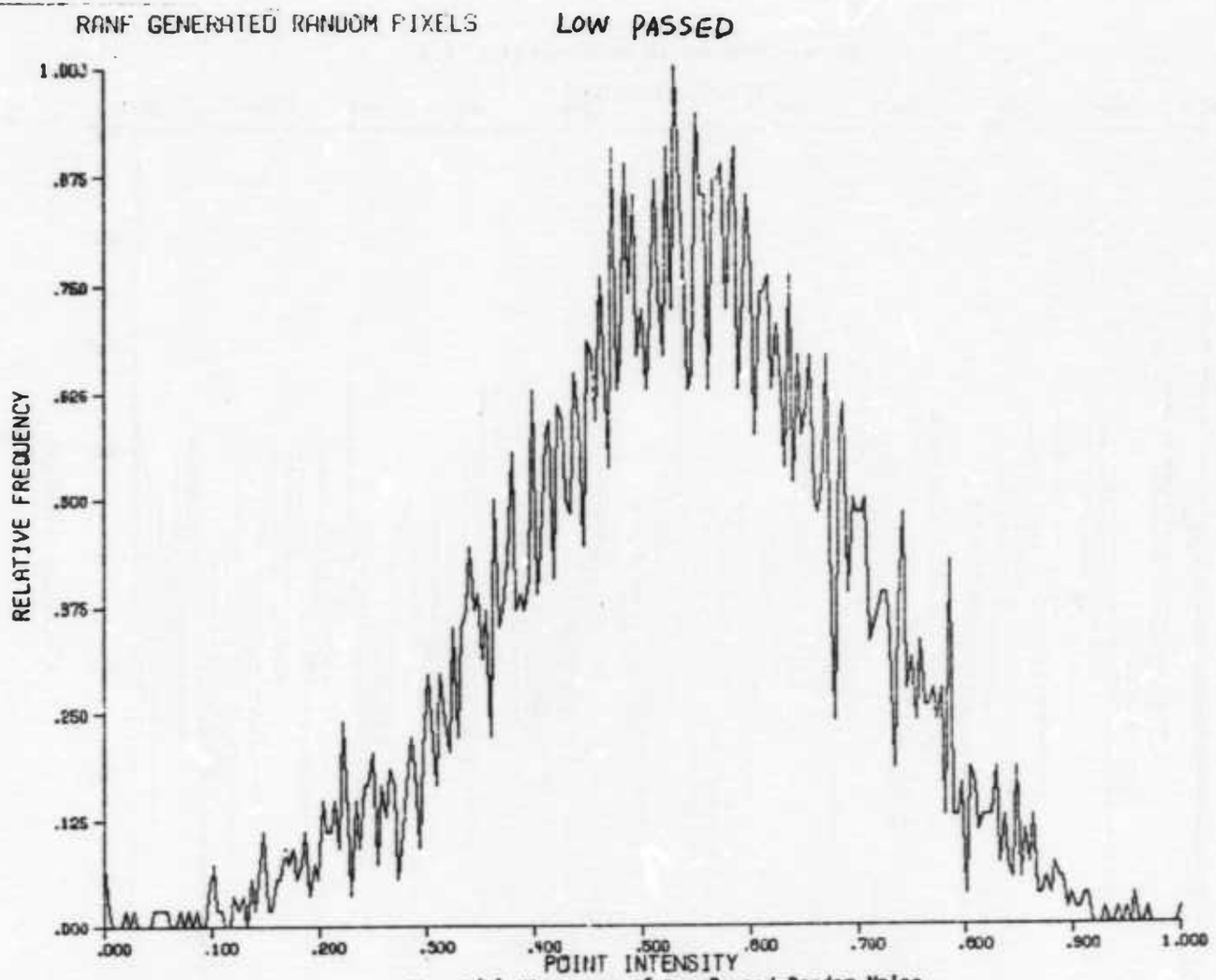


Fig. 2(b) Histogram of Low-Passed Random Noise
Mean = 0.53 Variance = .02 (64x64 Pixels)

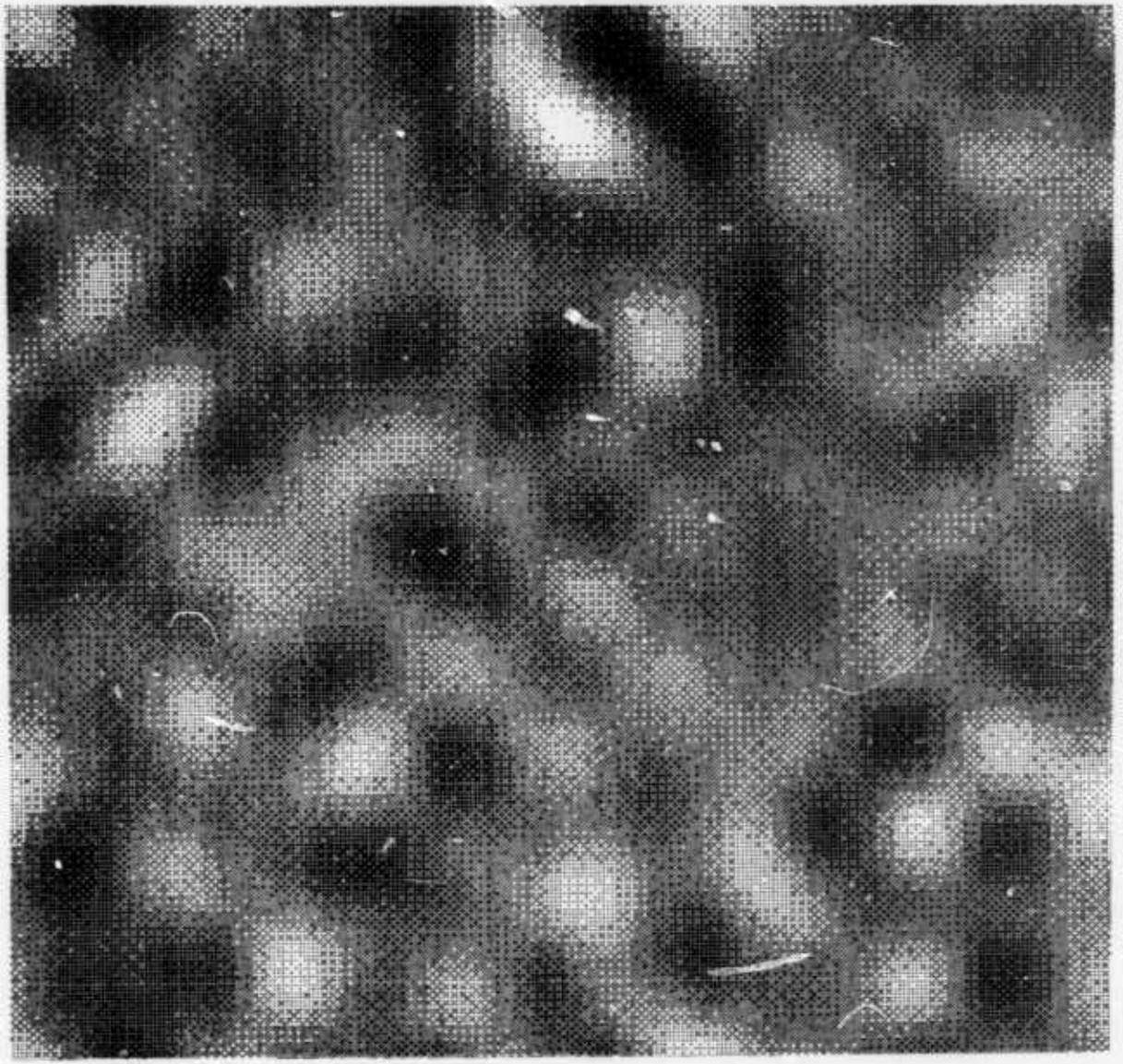


Fig. 2(c) Intensity Picture of Low-Passed Random Noise
Mean = .53 Variance = .02
(64x64 Pixels)

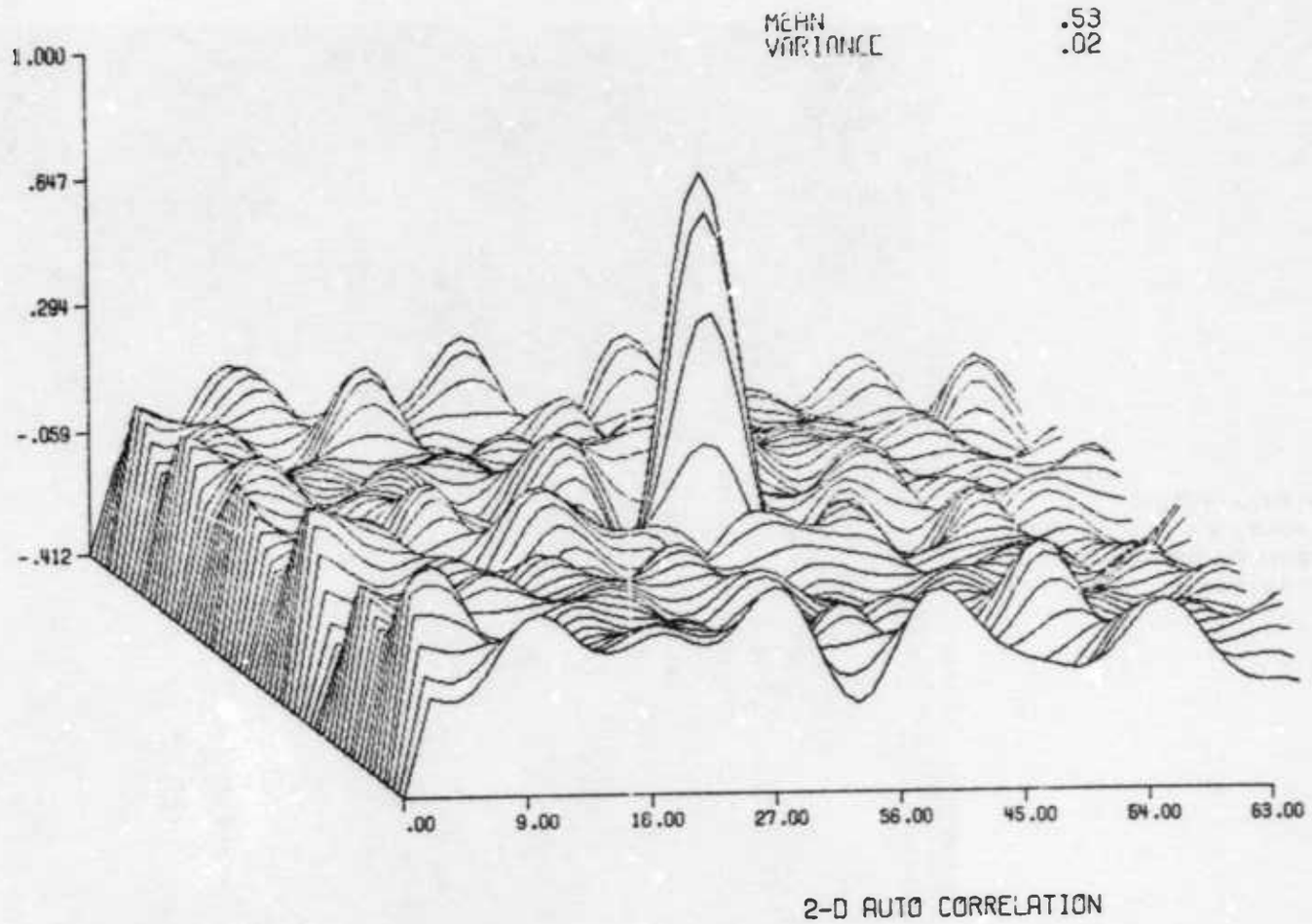


Fig. 2(d) Autocorrelation of the Low-Passed Random Noise

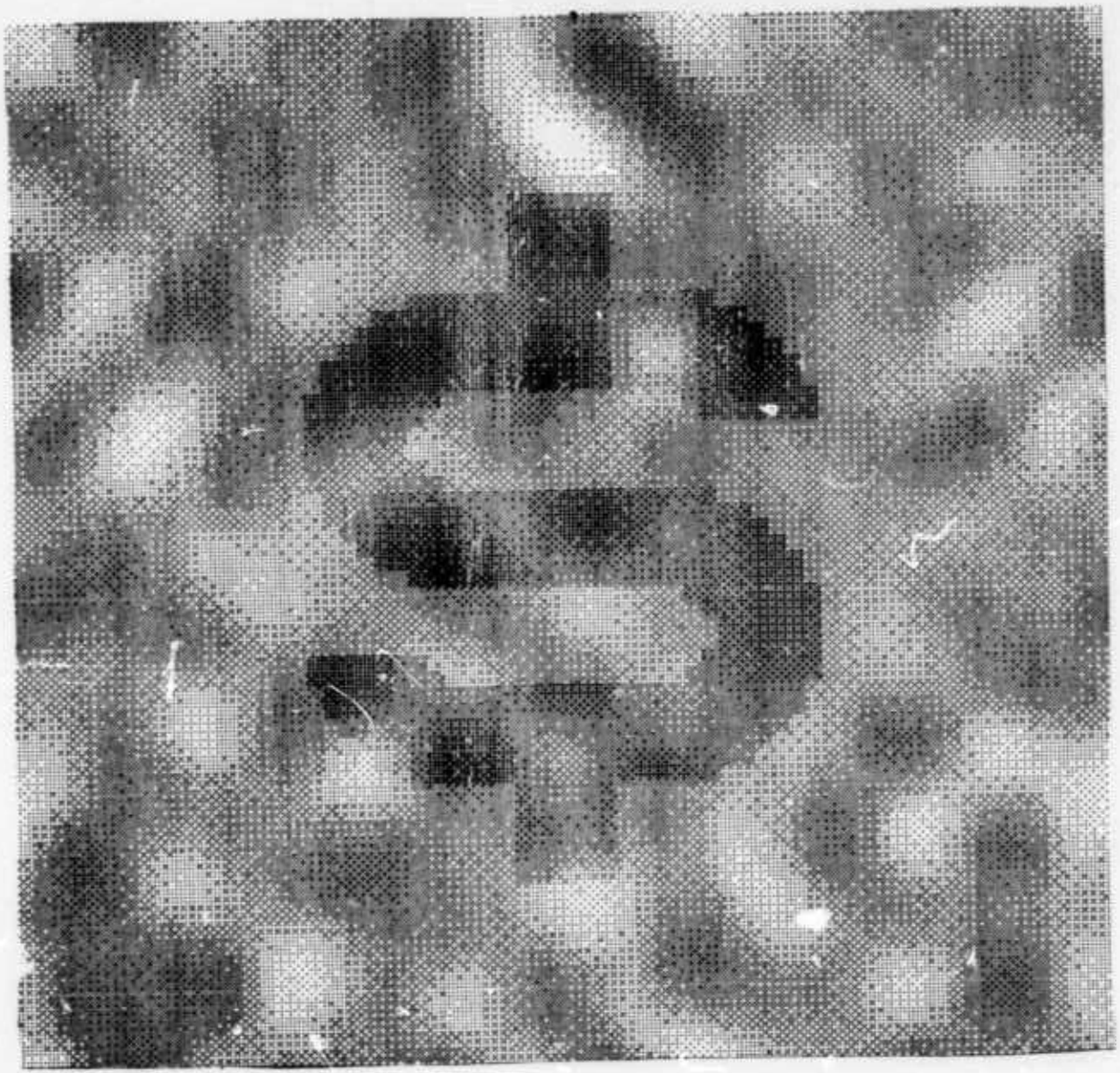


Fig. 3(a) Multiplicative Noise Picture with Noise Variance 0.02

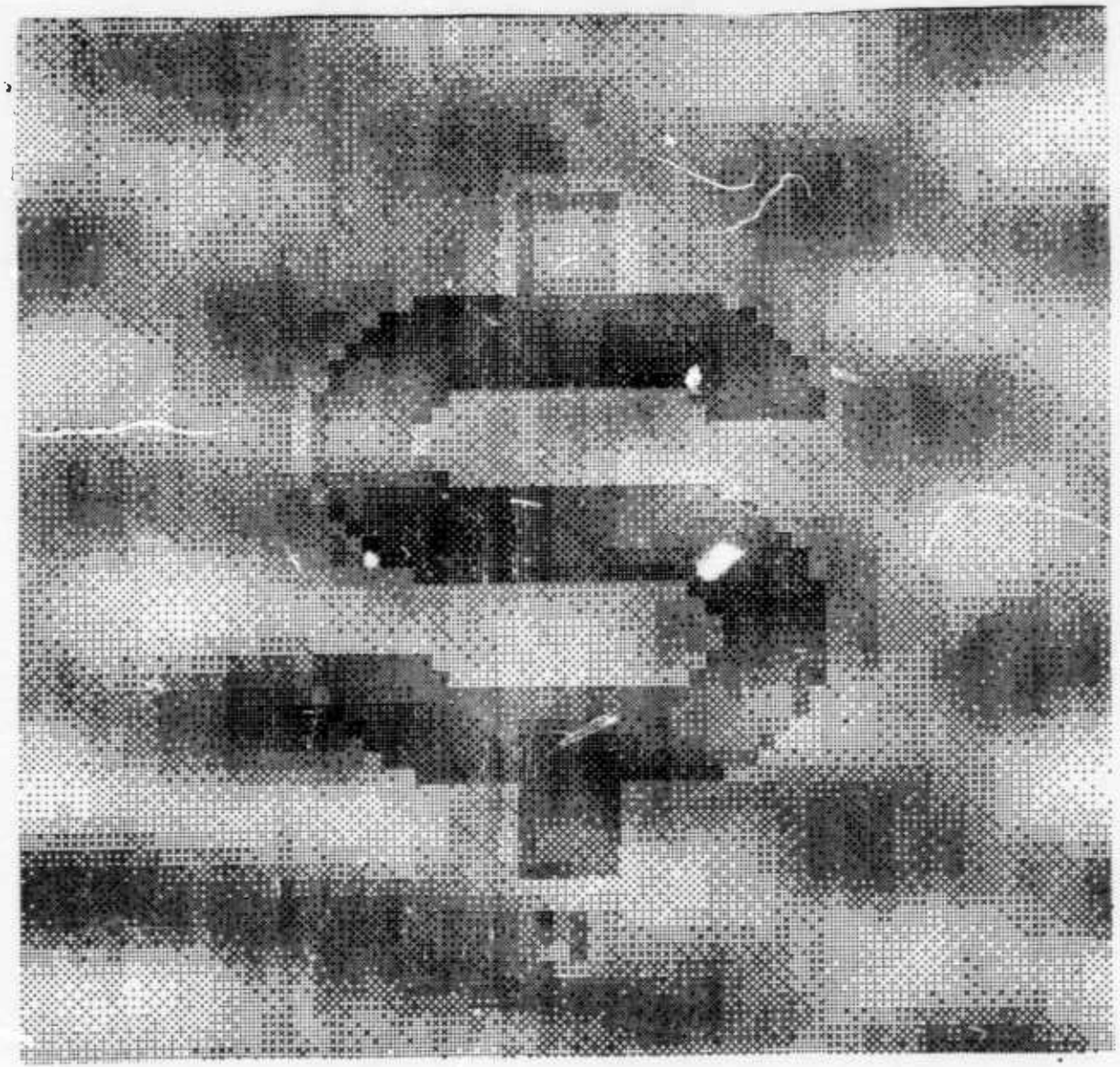


Fig. 3(b) The Filtered Picture of Fig. 3(a)

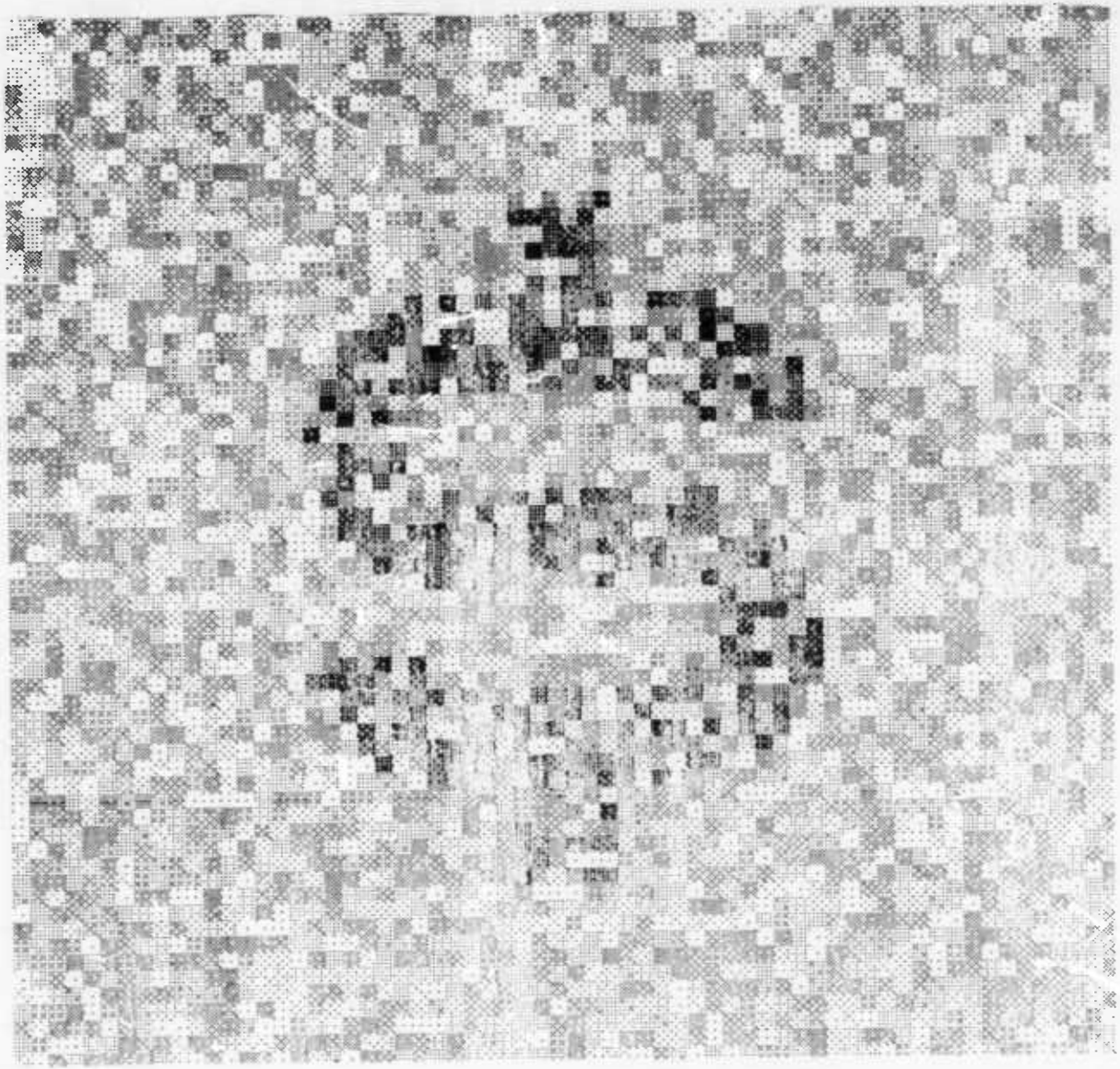


Fig. 3(c) Multiplicative White Random Noise Picture

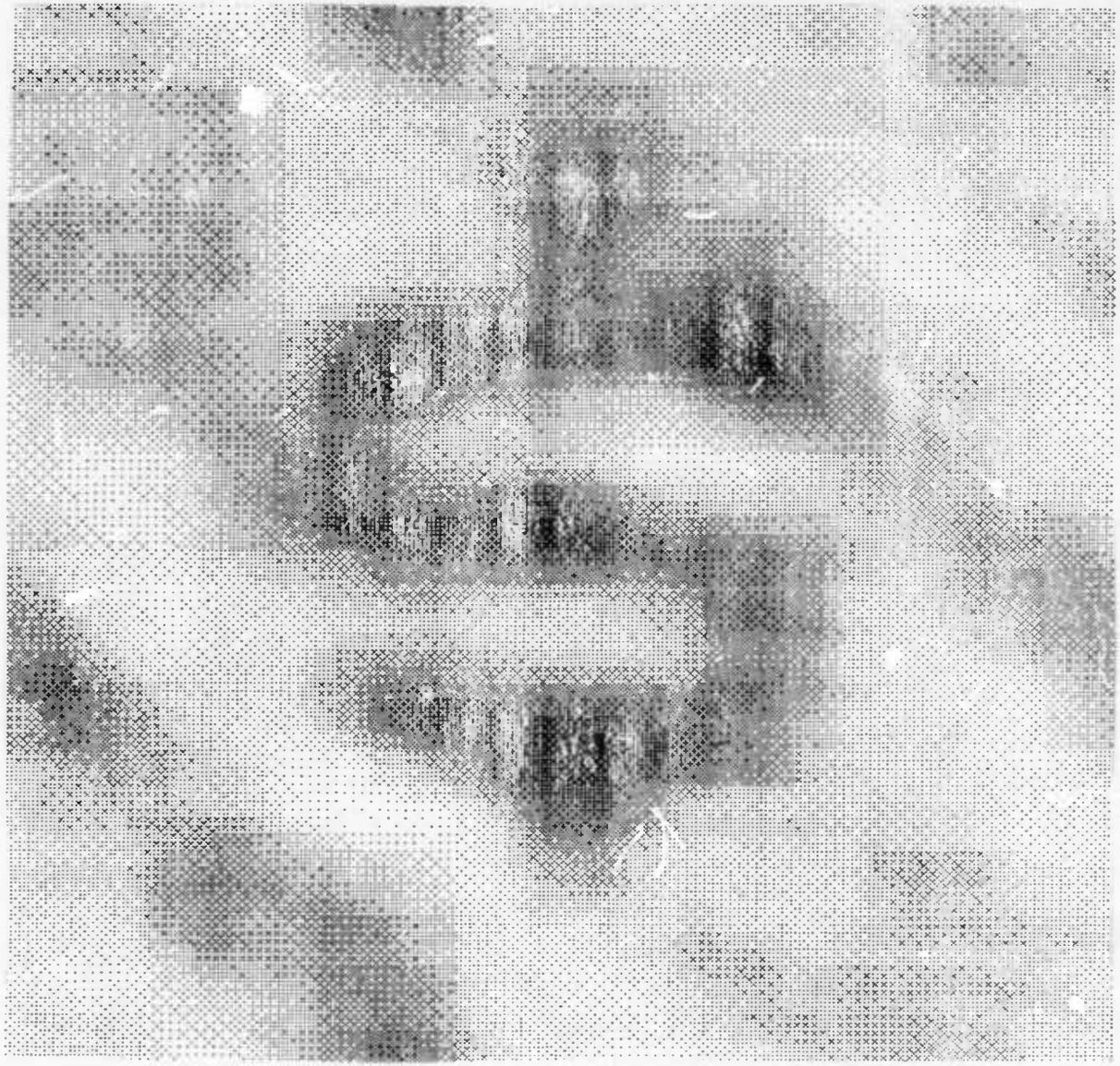


Fig. 3(d) The Filtered Picture of Fig. 3(c)

Figure 3(a) shows the original picture with multiplicative low-pass noise and Figure 3(b) shows the resulting filtered picture. Figure 3(c) shows the original picture with multiplicative white noise and Figure 3(d) is the resulting filtered picture.

A block diagram of the process is shown in Figure 4.

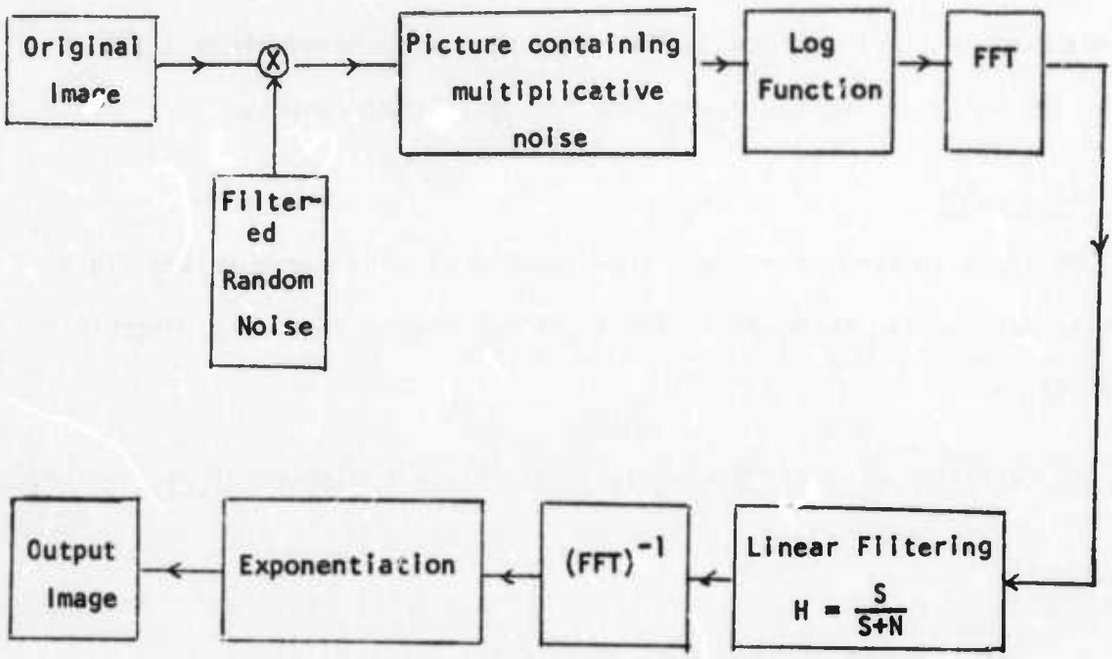


Fig. 4 Block diagram of homomorphic image filtering

We are now extending this model to ERTS imagery. The power spectrum of the signal and noise must be approximated since they are not directly available. The signal spectrum will be approximated by measuring clear ERTS frames; the noise spectrum will be obtained using multispectral channel classification on the noisy picture to produce a binary picture of the noise.

RECURSIVE RESTORATION OF LSI DEGRADED IMAGES

D. P. Panda and A. C. Kak

I. INTRODUCTION

A Kalman filter for restoration of images degraded by any LSI (linear shift invariant) point spread functions is developed here. Also mentioned here are certain flaws in some earlier work done by other authors [1,2] in developing Kalman filtering techniques for image enhancement.

II. IMAGE MODEL

The image is assumed to be a two-dimensional wide-sense Markov field [3] with exponential autocorrelation function, and defined on a discrete domain as

$$s(i,j) = \rho_v s(i-1,j) + \rho_d s(i-1,j-1) + \rho_h s(i,j-1) + \omega(i,j) \qquad (1)$$

where,

$s(i,j)$ = the undistorted image at the point (i,j)

ρ_v , ρ_h = vertical and horizontal correlation coefficient

$\rho_d = -\rho_v \rho_h$

$\omega(i,j)$ = A sequence of zero mean uncorrelated random variables.

The degraded observation $y'(i,j)$ is given by

$$y'(i,j) = \sum_{k'=-k'_0}^{k'_1} \sum_{\ell'=-\ell'_0}^{\ell'_1} s(i-k',j-\ell') \, g'(k',\ell') + n'(i,j) \qquad (2)$$

where

$g'(k',\ell')$ = the point spread function, assumed zero outside of

$$-k_0' \le k' \le k_1' \quad \text{and} \quad -\ell_0' \le \ell' \le \ell_1'$$

$n'(i,j)$ = A sequence of zero mean uncorrelated noise with variance σ_n^2

We shall now define

$$y(i,j) = y'(i-k_0', j-\ell_0')$$

$$p = k' + k_0'$$

$$q = \ell' + \ell_0'$$

$$k = k_1' + k_0'$$

$$\ell = \ell_1' + \ell_0'$$

$$n(i,j) = n'(i-k_0', j-\ell_0')$$

$$g(p,q) = g'(k',\ell') = g'(p-k_0', q-\ell_0')$$

then we get a shifted version of the observation

$$y(i,j) = \sum_{p=0}^{k} \sum_{q=0}^{\ell} s(i-p,j-q)\, g(p,q) + n(i,j) \qquad (3)$$

and,

$$E[n(i,j)n(k,\ell)] = \sigma_n^2\, \Delta(i-k)\Delta(j-\ell)$$

where, Δ = Kronecker delta.

If the size of the observed image $y(i,j)$ is mxn then we let

$$s(i) = [s(i,1)\ s(i,2)\ \ldots.\ s(i,n)]^T$$

$$y(i) = [y(i,1)\ y(i,2)\ \ldots.\ y(i,n)]^T$$

$$\omega(l) = [\omega(l+1,1)\ \omega(l+1,2)\ \ldots\ \omega(l+1,n)]^T$$

$$n(l) = [n(l,1)n(l,2)\ \ldots\ n(l,n)]^T$$

$$G(l) = \begin{bmatrix} g(l,0) & & & & \\ g(l,1) & g(l,0) & & & \\ \vdots & \ddots & \ddots & & 0 \\ g(l,\ell) & & g(l,1) & g(l,0) & \\ 0 & & g(l,\ell) & \ldots\ g(l,1) & g(l,0) \end{bmatrix}_{nxn}$$

$$R_v = \rho_v I_{nxn}$$

$$R = \begin{bmatrix} 1 & & & & \\ -\rho_h & 1 & & & \\ & -\rho_h & 1 & & 0 \\ & & \ddots & \ddots & \\ 0 & & & -\rho_h & 1 \end{bmatrix}_{nxn}$$

$$x(l) = [s^T(l)\ \ s^T(l-1)\ \ \ \ s^T(l-k)]^T$$

$$A = \left[\begin{array}{c|ccc} R_v & 0 & \cdots & 0 \\ \hline & & & 0 \\ I & & & \vdots \\ & & & 0 \end{array}\right]$$

$$B = [R^{-1}{}^T\ 0\ \ldots\ 0]^T$$

$$C = [G(0)\ G(1)\ \ldots\ G(k)]$$

The n we have

$$X(l+1) = A\,X(l) + B\,\omega(l)$$

$$y(l) = C\,X(l) + n(l) \qquad (4)$$

C = I (k=0 and ℓ = 0) gives the observation corrupted by additive noise only.

III. MOTIVATION FOR THE MODEL

As will be seen in later sections the model described in eqn. (4) will lead to a recursive restoration technique in which the estimate of the image $\hat{s}(i,j)$ is expressed in terms of past estimates and the observations along a complete row of the picture. Habibi and Nahi [1,2] have used the model described by eqn. (1) in enhancement of images corrupted by additive noise only, and have developed an estimate $\hat{s}(i,j)$ which is expressed in terms of previous estimates and the observation $y(i,j)$ only at the point (l,j). But, it can be shown that, in such a formulation the estimation error at the point (i,j), $e(i,j)$, is not necessarily, orthogonal to the observation $y(p,q)$ for $1 \leq p \leq i-1$, $q=j$ and for $p=i$, $1 \leq q \leq j-1$. In this sense the estimate is sub-optimal though computationally fast since it uses the observation only at one point, rather than on a complete row. It can also be shown that when this is extended to estimation of images degraded by LSI point spread function the suboptimality worsens. The recursive estimates are functions of estimates in the "past" and the observations at the "present". It was suspected that the above mentioned problem may be due to the fact that in a two-dimensional Markov field the "present" (as opposed to "past" and "future") is a curve [see e.g. 4,5] rather than just one point as in one-dimensional Markov sequences. The model described by eqn. (4) implicitly defines the "present" as one complete row rather than just one point and indeed this eliminates the suboptimality mentioned above.

IV. THE FILTER/PREDICTOR

We now can derive the recursive MMSE estimate of $X(i)$ in terms of estimates $\hat{X}(P)$, $i \leq P \leq i-1$, and $y(i)$ by following similar derivation for the estimate of a continuous time domain process as in [6]. If instead we wanted to estimate $X(i)$ in terms of $\hat{X}(P)$, $i \leq P \leq i-1$, and $y(i-1)$ we could use the result directly from [7]. Both the results are given below.

$$\hat{X}(i+1) = A\hat{X}(i) + H(i+1)[y(i+1) - CA\hat{X}(i)]$$

$$H(i+1) = Q(i)\, C^T[CQ(i)C^T + \Sigma_n]^{-1}$$

$$Q(i) = AP(i)A^T + B\Sigma\omega\, B^T$$

$$Q(i+1) = A[I - H(i+1)C]\, Q(i)A^T + B\Sigma\omega B^T \qquad (5)$$

And

$$\hat{X}(i+1) = A\hat{X}(i) + H(i)[y(i) - C\hat{X}(i)]$$

$$H(i) = AP(i)C^T[CP(i)C^T + \Sigma_n]^{-1}$$

$$P(i+1) = [A - H(i)C]\, P(i)[A - H(i)C]^T + B\Sigma\omega\, B^T + H(i)\Sigma_n H^T(i) \qquad (6)$$

where,

$$\Sigma_n = E[\omega(i)\omega^T(i)]\ ,\ \Sigma_n = E[n(i)n^T(i)]\ ,\ P(i) = E[e(i)e^T(i)]$$

Compare the result given in eqn. (5) with the corresponding result reported in [8].

The estimate of $s(i,j)$ is obtained by taking the jth row of $\hat{X}(i)$. It is necessary to have the initial condition $P(i)$ to determine $H(i)$, which is analogous to the coefficient $\eta(i,j)$ in the earlier work [9] reported by these authors. It is expected that just as with $\eta(i,j)$ in [9] the effect of initial condition on $H(i)$ will diminish as i increases. Also, because of modification in the definition of "present" the smearing of edges [see 9] is expected

to be less severe. In the presence of only LSI degradation and no additive noise H(i) and X(i) in eqn. (5) reduce to

$$H(i) = (C^T C)^{-1} C^T \tag{7}$$

and

$$\hat{X}(i) = H(i)y(i) \tag{8}$$

This gives the inverse filtering technique without going to the frequency domain and having to deal with zeroes in the transform of the point spread function. The sensitivity of the inverse filtering to any noise, however, is expected to be high.

It must be pointed out that the matrices C, A, Q, etc. can be very large even for a moderate size picture (e.g. 128x128) and point spread function (e.g. 10x10). In such a case the picture may be broken to several smaller size pictures concatenated and process each of the smaller pictures separately and concatenate the pieces back together. Similar work is being done for linear shift variant degradation of images.

V. THE PROOF OF EQN. (5)

Based on the model described by eqn. (4) we want the linear MMSE estimate $\hat{X}(i)$ of X(i) in terms of y(1), ..., y(i), i.e.,

$$\hat{X}(i) = \sum_{p=1}^{i} H(i,p)y(p) \tag{9}$$

with H(i,p) such that

$$E[\{X(i) - \hat{X}(i)\}y^T(q)] = 0 \qquad q \leq i$$

or

$$E[X(i)y^T(q)] = \sum_{p=1}^{i} H(i,p)E[y(p)y^T(q)] \tag{10}$$

$$E[X(i+1)y^T(q) = E[\{AX(i) + B\omega(i)\}y^T(q)]$$

$$= AE[X(i)y^T(q) + B\,E[\omega(i)y^T(q)$$

Since, $y(q) = CX(q) + n(q)$

$$= CAX(q-1) + CB\omega(q-1) + n(q)$$

$y(q)$, for all $q \leq i$, does not contain $\omega(i)$ and hence from eqn. (1)

$$E[\omega(i)y^T(q)] = 0$$

$$\therefore E[X(i+1)y^T(q)] = \sum_{p=1}^{i} AH(i,p)\,E[y(p)y^T(q)] \qquad (11)$$

But from eqn. (10),

$$E(X(i+1)y^T(q) = \sum_{p=1}^{i+1} H(i+1,p)\,E[y(p)y^T(q)]$$

$$= \sum_{p=1}^{i} H(i+1,p)E[y(p)y^T(q)] + H(i+1,i+1)E[y(i=1)y^T(q)]$$

$$E[y(i+1)y^T(q)] = E[\{CX(i+1) + n(i+1)\}y^T(q)]$$

$$= CE[X(i+1)y^T(q) + E[n(i+1)y^T(q)]$$

Since, for all $q \leq i$, $y(q)$ does not contain $n(i+1)$ $E[n(i+1)y^T(q)] = 0$.

Therefore,

$$E[y(i+1)y^T(q)] = \sum_{p=1}^{i} CAH(i,p)\,E[y(p)y^T(q)]$$

$$\therefore E(X(i+1)y^T(q)] = \sum_{p=1}^{i} [H(i+1,p) + H(i+1,i+1)CAH(i,p)E[y(p)y^T(q)] \qquad (12)$$

Combining eqn. (11) and (12)

$$\sum_{p=1}^{i} [(A - H(i+1,i+1)CA)H(i,p) - H(i+1,p)]E[y(p)y^T(q)] = 0$$

or, $$\sum_{p=1}^{i} H'(i,p)\, E[y(p)y^T(q)] = 0 \qquad (13)$$

where $$H'(i,p) = (A - H(i+1,i+1)CA)H(i,p) - H(i+1,p)$$

Let $$\hat{X}'(i) = \sum_{p=1}^{i} \{H(i,p) + H'(i,p)\} y(p)$$

$$\therefore E[\{X(i) - \hat{X}'(i)\}y^T(q)] = E[X(i)y^T(q)] - \sum_{p=1}^{i} H(i,p)E[y(p)y^T(q)]$$

$$- \sum_{p=1}^{i} H'(i,p)E[y(p)y^T(q)]$$

$$= 0 \text{ from eqn. (10) and (15)}$$

$\therefore$ Both $\hat{X}(i)$ and $\hat{X}'(i)$ are linear MMSE estimates of $X(i)$. Hence, from uniqueness of the linear MMSE estimate

$$\hat{X}(i) = \hat{X}'(i)$$

or $$\hat{X}(i) - \hat{X}'(i) = \sum_{p=1}^{i} H'(i,p)y(p) = 0 \qquad (14)$$

Since eqn. (14) is true for all values of Y(P) it must be true that

$$H'(i,p) = (A - H(i+1,i+1)CA)H(i,p) - H(i+1,p) = 0 \qquad (15)$$

$$\therefore \hat{X}(i+1) = \sum_{p=1}^{i} H(i+1,p)y(p)$$

$$= \sum_{p=1}^{i} (A - H(i+1,i+1)CA)H(i,p)y(p) + H(i+1,i+1)y(i+1)$$

$$= (I - H(i+1,i+1)C)A\hat{X}(i) + H(i+1,i+1)y(i+1)$$

$$= A\hat{X}(i) + H(i+1)[y(i+1) - CA\hat{X}(i)] \tag{16}$$

by writing $H(i+1)$ for $H(i+1,i+1)$.

Hence, the error $e(i+1) = X(i+1) - \hat{X}(i+1)$

$$= AX(i) + B\omega(i) - [A\hat{X}(i)H(i+1)\{CAX(i) + CB\omega(i) + n(i+1) - CA\hat{X}(i)\}]$$

$$= (I - H(i+1)CAe(i) + (I - H(i+1)C)B\omega(i) + H(i+1)n(i+1) \tag{17}$$

Since, $E[e(i+1)y^T(q)] = 0$ for all $q \leq i+1$

$$E[e(i+1)\hat{X}^T(i)] = 0$$

$$\therefore E[e(i+1)\{y(i+1) - CA\hat{X}(i)\}^T] = 0$$

or $$E[e(i+1)\{CAe(i) + CB\omega(i) + n(i+1)\}^T] = 0 \tag{18}$$

Since $$E[\omega(i)y^T(q)] = 0 \quad q \leq i$$

$$E[\omega(i)e^T(i)] = 0$$

Substituting (17) in (18) and writing $E[e(i)e^T(i)] = P(i)$ we get,

$$(I - H(i+1)C)AP(i)A^TC^T + (I - H(i+1)C)B\Sigma_\omega B^TC^T + H(i+1)\Sigma_n = 0$$

or $$H(i+1) = (AP(i)A^T + B\Sigma_\omega B^T)C^T[C(AP(i)A^T + B\Sigma_\omega B^T)C^T + \Sigma_n]^{-1}$$

$$= QC^T[CQC^T + \Sigma_n]^{-1} \tag{19}$$

$$P(i+1) = E[e(i+1)e^T(i+1)]$$

$$= E[e(i+1)X^T(i+1)]$$

$$= E[e(i+1)\{AX(i) + B\omega(i)\}^T]$$

Substituting eqn. (17) we have

$$P(i+1) = \{ I - H(i+1)C \} AP(i)A^T + \{ I - H(i+1)C \} B\Sigma_\omega B^T$$

$$= [I - H(i+1)C] Q(i)$$

or $$Q(i+1) = AP(i+1)A^T + B\Sigma_\omega B^T$$

$$= A[I - H(i+1)C]Q(i)A^T + B\Sigma_\omega B^T \qquad (20)$$

Putting together eqn. (16), (19), and (20) we get eqn. (5).

REFERENCES

[1] A. Habibi, "Two-Dimensional Bayesian Estimate of Images," Proc. IEEE, v. 60, N. 7, pp. 878-883, July 1972.

[2] N. E. Nahi and A. Habibi, "Decision-Directed Recursive Image Enhancement," IEEE Trans. on Circuits and Systems, vol. CAS-22, No. 3, pp. 286-293, March 1975.

[3] A. Rosenfeld and A. C. Kak, Digital Picture Processing, Academic Press, New York, to be published.

[4] J. W. Woods, "Two-Dimensional Discrete Markovian Fields," IEEE Trans. on Inf. Th., Vol. IT-18, No. 2, pp. 232-240, March 1972.

[5] E. Wong, "Recursive Filtering for Two-Dimensional Random Fields," IEEE Trans. on Inf. Th. pp. 84-86, January 1975.

[6] R. E. Kalman and R. S. Bucy, "New Results in Linear Filtering and Prediction Theory," Trans. of the ASME, J. Basic Eng., pp. 95-108, March 1961.

[7] N. E. Nahi, Estimation Theory and Application, John Wiley, 1969.

[8] A. O. Aboutalib and L. M. Silverman, "Restoration of Motion Degraded Images," IEEE Trans. on Circuts and Systems, Vol. CAS-22, No.3, pp. 278-286, March 1975.

[9] A. C. Kak and D. Panda, "Image Enhancement by Recursive Estimation," in Image Analysis and Modelling, Annual Research Report under ARPA, TR-EE 75-2, School of Electrical Engineering, Purdue University, October 1, 1973-October 31, 1974.

CHARACTERIZING THE INFORMATION CONTENT OF MULTISPECTRAL IMAGES

P. A. Wintz

I. INTRODUCTION

The information contained in multispectral imagery of the earth's surface is spread through four dimensions - 2 spatial, 1 spectral, and 1 temporal. Whether the information is extracted by a person or a machine the magnitude of the information extraction task could be lessened if the data bulk could first be reduced. In the case of machine processing there are three sources of savings - the time required to input the data into the machine, the amount of memory required to store the data, and the amount of machine time required to extract the information.

The problem is to minimize the cost of extracting information from the data under the constraint that the degradation in the performance of the information extractor, e.g., classification accuracy, not exceed a prescribed amount.

Some source coding algorithms may be reasonably efficient for our purpose. An algorithm that appears to serve both purposes in some applications is the principal components (eigenvector) transformation. Fig. 1 shows the results on an experiment performed on some 12 channel aircraft data. The right hand scale gives the mean square error between the original and reconstructed images and the left hand scale lists the percent classification accuracy obtained by inputing the transformation coefficients directly into the pattern recognition machine. Fig. 1a resulted from using the transforming 10x10 sub arrays in the spatial domain. Compression ratios of 4 in the spectral domain and 10 in the spatial domain were achieved before classification accuracy degraded. Since the number of computations required to implement the maximum likelihood algorithm for Gaussian data is proportional

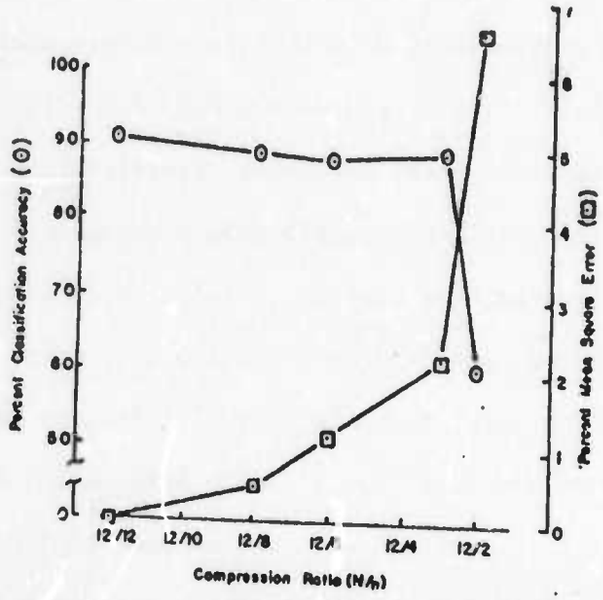


Fig. 1a Spectral Compression

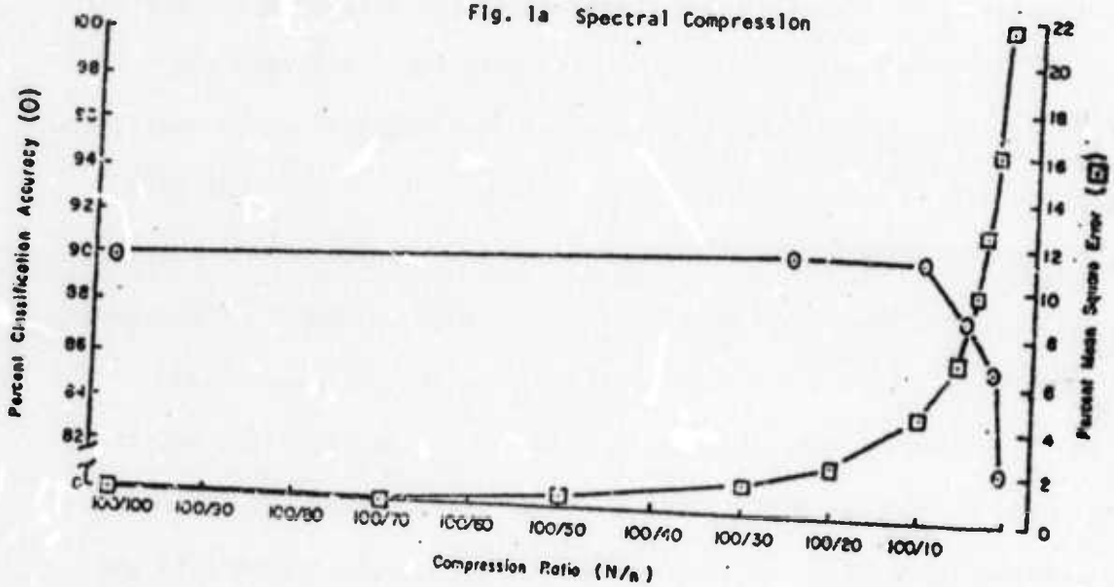
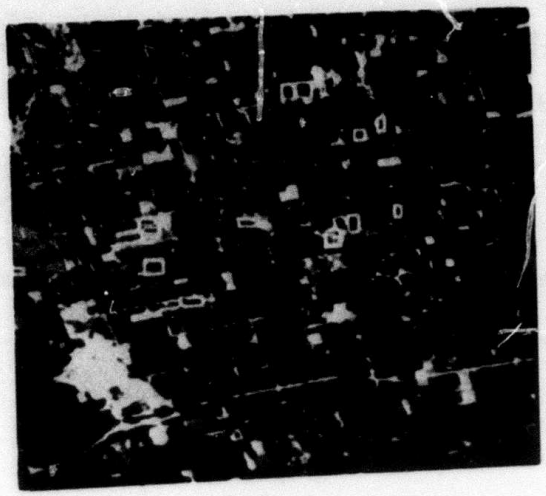


Fig. 1b Spatial Compression

to the square of the dimensionality of the data, a significant reduction in the computational load of the pattern recognition machine was achieved.

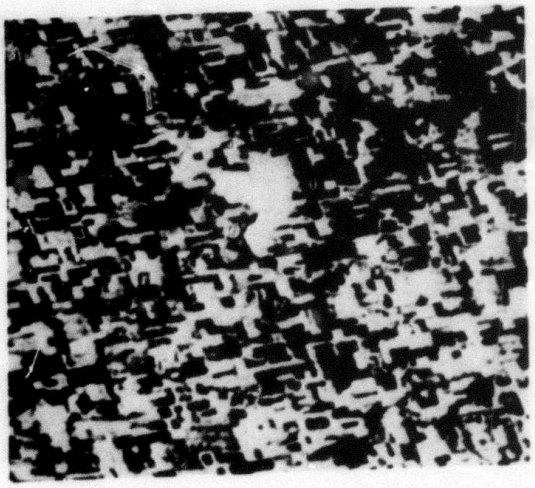
Any solution to this problem must be both data and data user dependent. The data dependence is similar to the data dependence of source codes: they must be matched to the characteristics, structure, statistics, etc. of the data. The dependence on the data user and his data requirements can be illustrated by looking at two extreme cases: If the user data requirements are not well defined (we do not know a priori who the data users will be or what they will require of the data) then the only satisfactory solution may be to allow no distortion of the data, i.e., use an information preserving code. At the other end of the scale suppose that we know that all potential users will do spectral signature analysis and that we also know the classifications. Then we could classify each resolution element and code the classification map. If K classes are required then $\log_2 K$ bits per resolution element are needed to code the classification map. If the original data consisted of N multispectral channel and M bits per resolution element then the original data set contained NM bits per resolution element. For example ERTS data has NM=28 whereas a classification map with 32 classes would require $\log_2 32=5$ bits per resolution element. The classification map could itself be coded by contour coding, the Duan-Wintz [4] folded DPCM method, etc., to achieve fewer bits but the cost of decoding would probably be prohibitive.

Consider an intermediate case. Suppose that the set of users are constrained to spectral signature analysis, but the classification categories are unknown. Can we provide the data user(s) with a significantly reduced set of data in a format that minimizes the amount of computation he must do while allowing arbitrary classification categories with negligible degradation in classification accuracy?

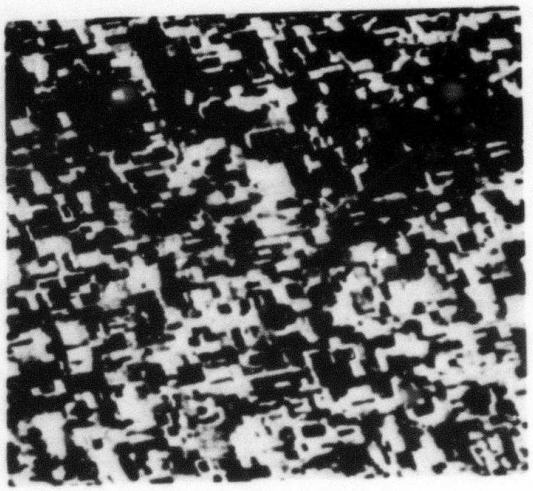


(a) Channel 1

(b) Channel 2



(c) Channel 3



(d) Channel 4

Fig. 2 Sections of a frame of ERTS-1 MSS Data
(LARS Run No. 72032806)

II. ADAPTIVE FEATURE SPACE CODING

Our approach to this problem is to determine all the distinguishable clusters in the spectral dimension of the data and to code the clusters. That is, we first use a clustering algorithm to determine the clusters; then we determine the mode center of each cluster; then we label each mode center; finally, we generate a cluster map, i.e., a map with each resolution element labelled with its cluster number. Since every conceivable category corresponds to one of the clusters, any data user can classify the data into any set of categories using a table look up procedure. That is, he first determines which set of clusters correspond to which category (training) to generate the table, and then he classifies each resolution element by using its cluster label to address the table. For the ERTS frame presented in Fig. 2 we found 96 distinguishable spectral domain clusters so that the cluster map required 7 bits per resolution element whereas the original ERTS data required 28 bits per resolution element. Every conceivable classification category corresponds to one or more of the clusters so that any data user can classify any resolution element by mapping its cluster label into the appropriate category label.

As an example, we used this method on data from channels 2 and 3 of the ERTS frame presented in Fig. 2. With a window of 6x6 resolution elements around each mode center in the 2-dimensional feature space our algorithm found 96 distinguishable clusters so that the cluster map required 7 bits per resolution element. Applying the Duan-Wintz folded DPCM algorithm to the cluster map reduces this to 4 bits per resolution element. With this cluster map as the input data to the classification algorithm with the three classes corn, soybeans, and others we achieved a classification accuracy of 77%. The standard LARSYS maximum likelihood algorithm operating on the original data achieved a classification accuracy of 78% for the same classes.

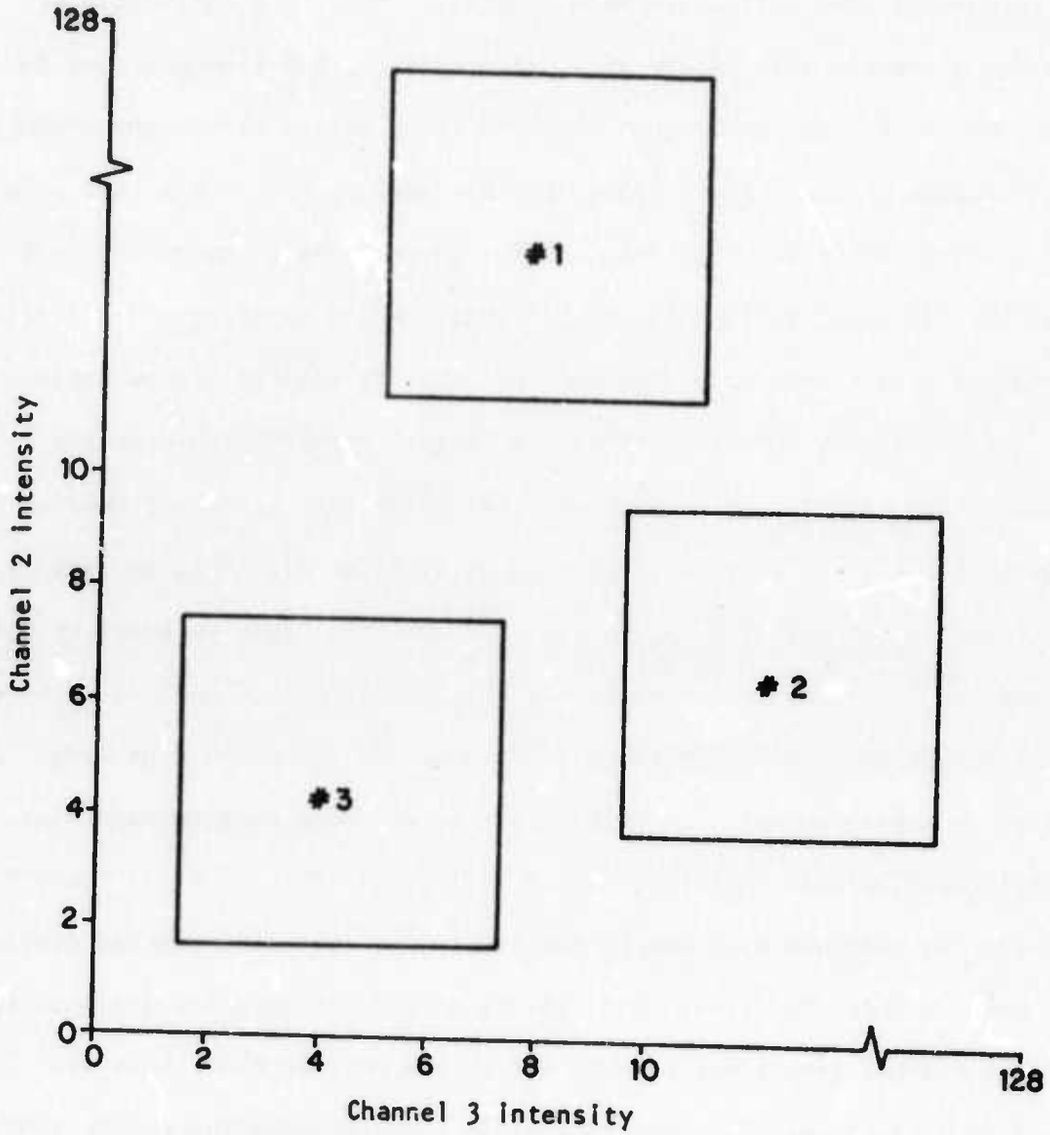


Fig. 3 The first three 6x6 clusters (adaptive 2-dimensional quantization

Let us now look at this method and, in particular, the example of the preceding paragraph from a different point of view. The 2-dimensional feature space obtained by plotting the channel 2 intensity against the channel 3 intensity consists of 127x127 points and each ERTS resolution element falls into one of these points. Our method achieves a data bulk reduction by requantizing this space more coarsely. This is accomplished by first making a scatter plot of the data, then moving a 6x6 window around to find the location having the largest number of data points within the window. This 6x6 subspace of the feature space is then labelled #1 and all data points within the window are given this label. This corresponds to constructing a 2-dimensional 6x6 quantization bit at this location and rounding off all data within the bin to a common quantization level. Next, the 6x6 window again moved to again find the location having the largest number of data points within the window only this time with the constraint that it cannot overlap quantization bin #2. This second 6x6 array is labelled #2. This process is repeated until we run out of data points. For the ERTS frame we ended up with 96 bins ordered according to the number of data points within each bin. Some regions of the feature space contained no data points and so no bins were constructed in these regions. In general, there are some gaps between bins that contain data points. Refer to Fig. 3 and suppose that we first found bin #1, then bin #2, then bin #3. Due to the constraint that bins can not overlap gaps can occur between the bins. Data points within the gaps are assigned the label of the nearest bin, i.e., rounded off to the nearest bin. This is equivalent to letting the bins grow to take in these neighboring points so that, finally, some of the quantization bins are irregular in shape (not 6x6). For the data of Fig. 2 this 2-dimensional coarse requantization of the data led to a 1% degradation in the classification accuracy.

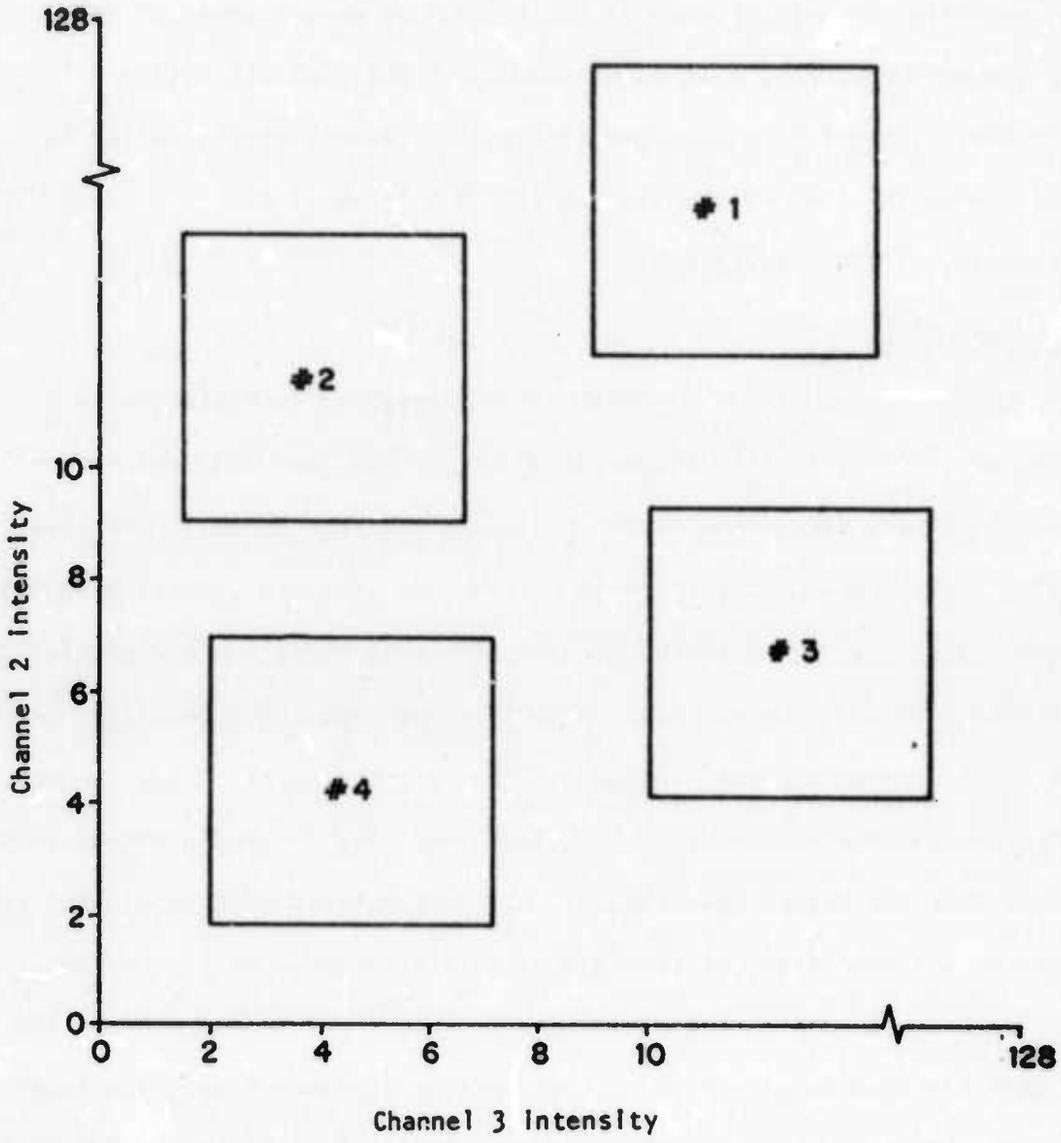


Fig. 4 Nonadaptive 2-dimensional quantization bins

III. NONADAPTIVE FEATURE SPACE CODING

Another way to more coarsely quantize the 2-dimensional feature space is to fill the feature space with a regular array of 6x6 quantization bins as illustrated in Fig. 4. This corresponds to more coarsely quantizing both of the channel intensities and would correspond to returning fewer significant bits for each for the natural code if the array size were a power of two. Whereas, the method of Fig. 3 is data dependent (adaptive) the method of Fig. 4 is not. Obviously, the method of Fig. 4 is significantly easier to implement. When this method was used on the ERTS frame of Fig. 2 a classification accuracy of 73% was achieved.

IV. STAGGERED DPCM

One method for reducing the number of bits required to characterize a multispectral image is to take advantage of the spatial and spectral redundancy by simply neglecting some of the pels. In the reconstruction process the missing picture elements are estimated from the retained picture elements and inserted into the reconstructed image.. The retained picture elements are coded by DPCM. To minimize the loss in spatial and spectral resolution the retained picture elements from successive lines are staggered, e.g., if the even numbered pels are deleted from the first row, then the odd numbered pels are deleted from the second row, etc. so that each deleted picture element is surrounded on all four sides by retained picture elements and, in the reconstruction process, can be estimated from its four nearest neighbors. The deleted pels can also be staggered in the spectral dimensions as illustrated in Fig. 5 so that half the data points in each spectral vector are retained.' A range of data compression ratios can be obtained by varying the number of bits, i.e., the number of quantization levels, used to code the pel differences.

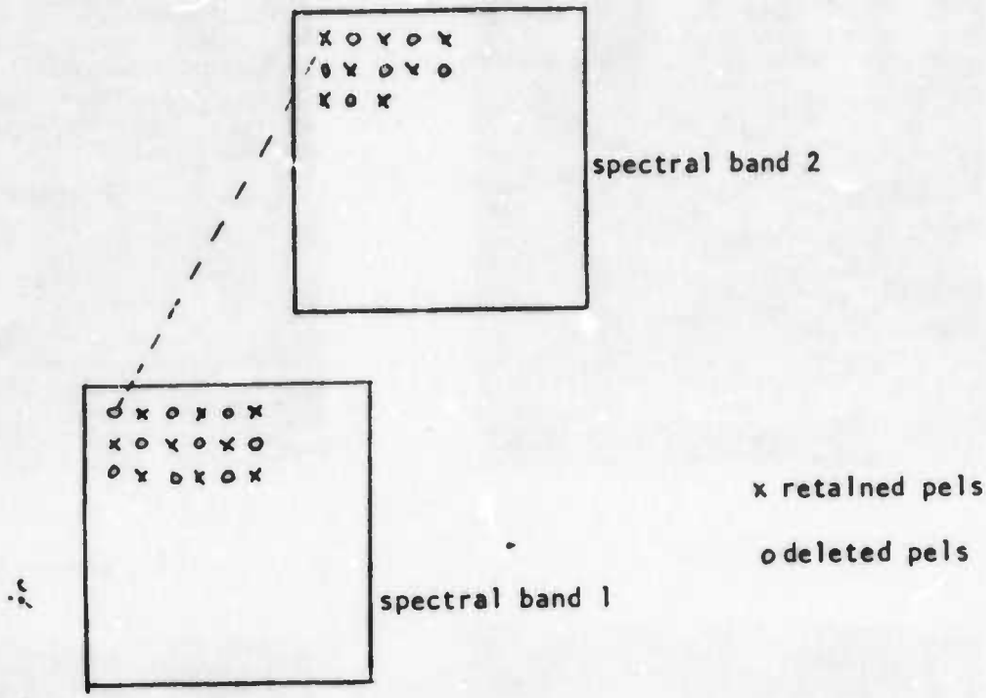
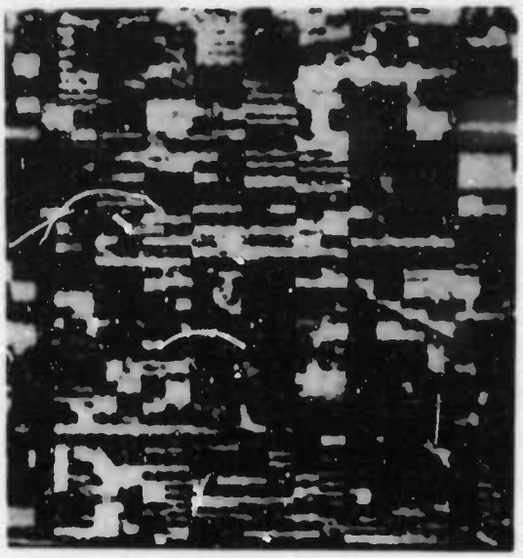


Fig. 5

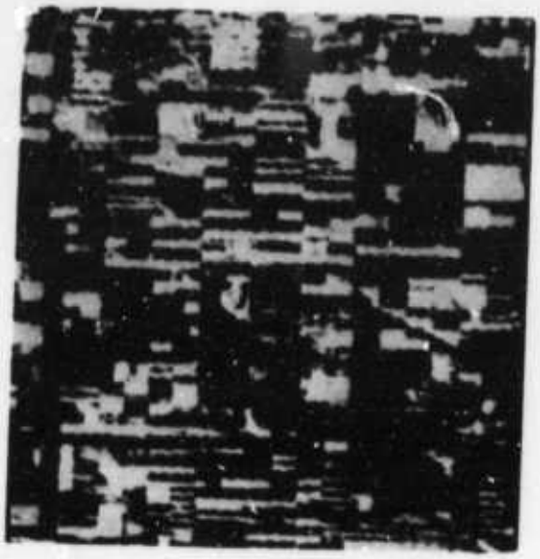
This technique was tested on two ERTS frames, LARS run no. 73041801 presented in Fig. 2 and LARS run no. 72032806 presented in Fig. 6. The reconstructed data was inputed to the standard LARSYS maximum likelihood classification algorithm and the resulting classification accuracies are presented in Fig. 7 for a range of compression ratios. The degradation in classification accuracy relative to the accuracy obtained from the original data (which corresponds to a compression ratio of 1 in Fig. 7) is surprisingly small out to compression ratios of 7.

We also generalized the same basic strategy to retain every third pel and the performance of this strategy is also plotted in Fig. 7 at a compression ratio of 9. It appears that this does not work as well as retaining every second pel and using a coarser quantizer to achieve the same compression ratio.

Also plotted on Fig. 7 are the performances of the two cluster coding techniques discussed in the preceding sections.



(a) Channel 1



(b) Channel 2

(c) Channel 3

(d) Channel 4

Fig. 6 Sections of a frame of ERTS - 1 MSS Data (LARS Run No. 73044801)

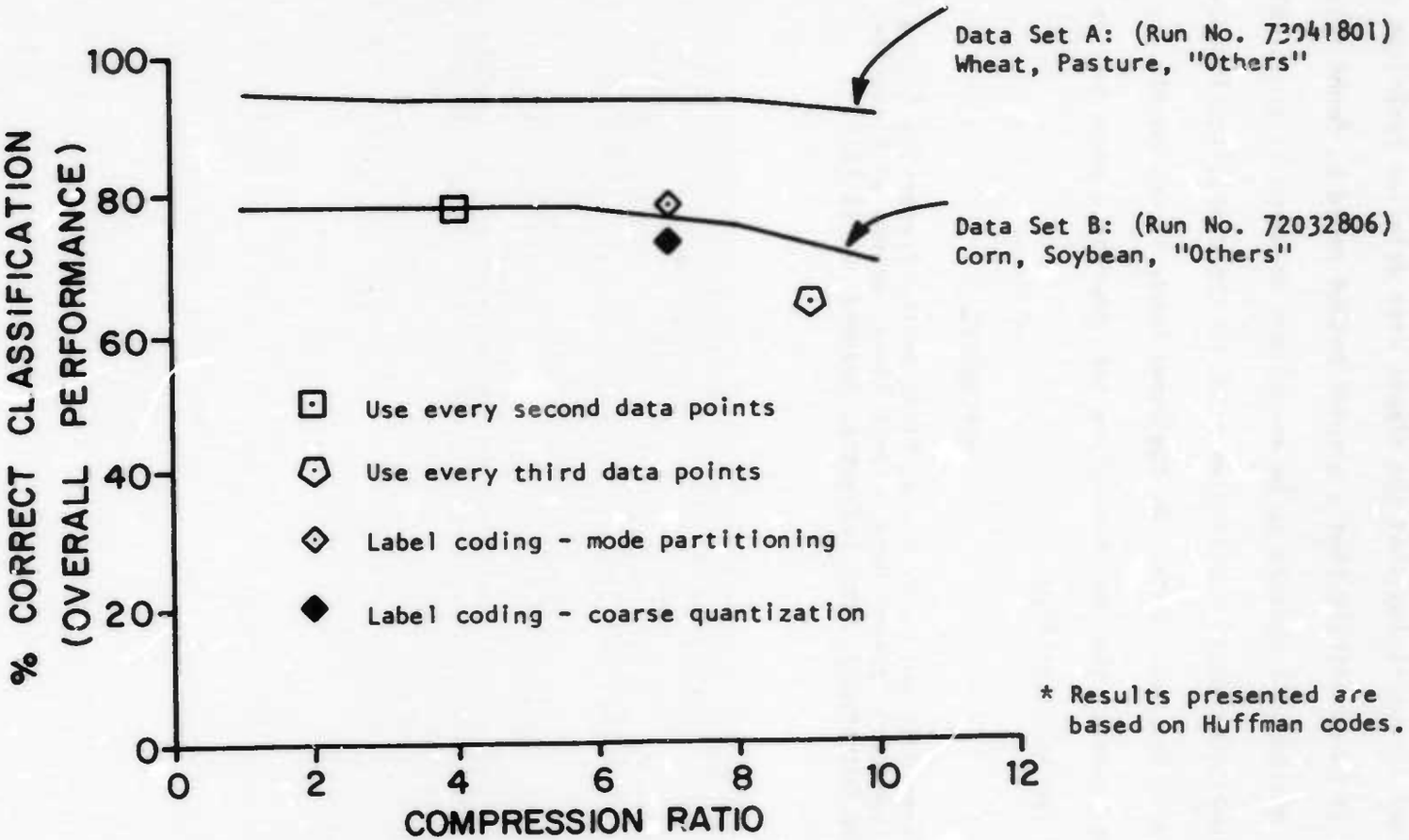


Fig. 7 Classification Accuracy vs. Compression Ratio*
For data compressed by using staggered DPCM and two other schemes as indicated.

We find it surprising that the simple data deletion technique performs so well. It is not likely that a simpler method could be found that performs this well and so this appears to be an optimum solution to the problem of image coding relative to a criterion based on the classification accuracy achieved with the coded data. On the other hand, it has no utility to the problem we posed at the beginning since the compressed data must be decoded before it can be classified.

REFERENCES

[1] Jen-Reen Duan and P. A. Winta, "Information-Preserving Coding for Multispectral Scanner Data," TR-EE 74-15, School of Electrical Engineering, Purdue University, West Lafayette, Indiana, April 1974.

FACILITIES

Present local facilities available to us include:

1 PDP 11/45 computer; 32K core; disk operating system software

1 20-million word disk

2 9-track tape drives; 800 and 1600 bpi densities

1 7-track tape drive; 556 and 800 bpi densities

3 CRT terminals

1 card reader

1 paper tape reader

1 line printer; 600 lines per minute, impact type

1 electrostatic graphics hard copy system, 200 points per inch

1 video display with disk storage; 3 color refreshed display, 512x512 points, 8 bits/color/point; graphics overlay and interactive capabilities

1 flying spot scanner

The hardware interface between our computer and the IMP has been completed. We are currently developing an ELF software system for use on our system. We are also investigating the use of UNIX as a replacement for both DOS and ELF.

Basic network data transfer capabilities will be available in June 1975 and complete hardware and software capabilities will be available in late 1975.

In addition to our own facilities we have the use of the Purdue University Computing Center's CDC 6500 and the IBM 360/67 at the Laboratory for Application of Remote Sensing. Batch picture processing programs with graphic output have been written for the CDC 6500, and the LARSYS operating system at LARS allows us to do interactive processing of multispectral data. Local interactive processing is also available on our PDP 11/45 computer. A software package (AMIDS) developed at RADC for pattern analysis is being transferred to our PDP.

JOURNAL PUBLICATIONS

HUANG, T. S. and KOLLER, H. U., "Coding of Multi-level Graphics," to appear in IEEE Trans. on Communications.

HUANG, T. S., BARKER, D., and BERGER, S., "Iterative Image Restoration," to appear in Applied Optics.

HUANG, T. S., "Bounds on Entropies of Linear Codes," to appear in IEEE Trans. on Inf. Theo.

HUANG, T. S. and NARENDRA, P., "Image Restoration by Singular Value Decomposition," to appear in Applied Optics.

HUANG, T. S. and HUSSIAN, A. B. S., "Facsimile Coding by Skipping White," to appear in IEEE Trans. on Communications.

HUANG, T. S., BURNETT, J., and DECZKY, A., "The Importance of Phase in Image Processing Filters," to appear in IEEE Trans. on Audio, Speech, and Signal Processing.

KAK, A. C., BEAUMONT, L. R., and WOLFLEY, J., "Signal Processing in the Determination of Acoustic Impedance Profiles," TR-EE 75-7, School of Electrical Engineering, Purdue University, West Lafayette, IN 47907.

PANDA, D. P., PATRICK, E. A., STELMACK, F., and JARDINA, S., "Pattern Recognition Applied to Surgery," Computers in Biology and Medicine, Vol. 4, 1975.

WINTZ, P. A. and GUPTA, J., "A Boundary Finding Algorithm and its Applications," IEEE Trans. on Circuits and Systems, March 1975.

WINTZ, P. A. and WILKINS, L. C., "On the Effect of Timing Errors in Run Length Codes," IEEE Trans. on Communications, June 1975.

FUKUNAGA, K., "The Estimation of the Gradient of a Density Function and its Applications in Pattern Recognition," to appear in IEEE Trans. on Information Theory, (with L. D. Hostetler).

FUKUNAGA, K., "k-Nearest Neighbor Bayes Risk Estimation," to appear in IEEE Trans. on Information Theory, (with L. D. Hostetler).

FUKUNAGA, K., "A k-Nearest Neighbor Mean-Square-Error Approach to Bayes Risk Estimation," to appear in IEEE Trans. on Information Theory, (with L. D. Hostetler).

FUKUNAGA, K., "Unbiased Estimation of Polynomial Functions Involving Probability Densities," to appear in IEEE Trans. on Information Theory, (with L. D. Hostetler).

FUKUNAGA, K., "Composite Goodness of Fit Tests for Distribution with Complete Sufficient Statistics," to appear in IEEE Trans. on Information Theory (with D. L. Kessell).

FUKUNAGA, K., "A Test for Multivariate Normality with Unspecified Parameters," to appear in IEEE Trans. on Information Theory, (with D. L. Kessell).

FUKUNAGA, K., "Pseudo Eigenvectors and their Applications to Pattern Recognition," to appear in IEEE Trans. on Electronic Computers, (with G. V. Sherman).

FUKUNAGA, K., "A Branch and Bound Clustering Algorithm," to appear in IEEE Trans. on Computers, (with P. M. Narendra).

FUKUNAGA, K., "A Branch and Bound Algorithm for Computing the k-Nearest Neighbors," to appear in IEEE Trans. on Computers, (with P. M. Narendra).

CONFERENCES

HUANG, T. S. and TANG, G. Y., "Application of Edge Detection to Image Enhancement," presented at the 5th Annual Symposium on Automatic Imagery Pattern Recognition, University of Maryland, Silver Springs, MD, April 17-18, 1975.

HUANG, T. S., Panel Discussion on Optical/Digital Signal Processing, (Invited), International Computing Conference, Washington, D. C., April 23-25, 1975.

HUANG, T. S., "Electronic Processing in Flaw Detection," (Invited), presented at the SPIE Symposium on the Use of Optics to Solve Reliability and Quality Control Problems, San Diego, CA, May 15-16, 1975.

HUANG, T. S., "Some Easily Implementable Suboptimum Runlength Codes," (Invited), to be presented at the 1975 International Communications Conference, San Francisco, CA, June 16-18, 1975.

HUANG, T. S., "Coding of Two-tone Image," (Invited), to be presented at the SPIE Seminar on Image Coding, San Diego, CA, August 21-22, 1975.

MITCHELL, O. R. and MYERS, C. R., "Statistical Measures for the Determination of Texture Primitives," presented at the Fifth Annual Symposium on Imagery Pattern Recognition, College Park, MD, April 17-18, 1975.

MITCHELL, O. R. and JONES, W. P., "Information Extraction Techniques for Drug Related Effects on the Evoked EEG," presented at the Midwest Biomedical Engineering Conference and Workshop, Columbus, OH, April 11-12, 1975.

WINTZ, P. A., WILSON, J., and SCHMIDT, E., "TV Bandwidth Compression," (Invited), Proc. of the SPIE 19th Annual Technical Symposium, August 1975.

WINTZ, P. A., "Image Filtering, Coding, and Information Extraction," Seminar presented at Stanford University, Stanford, CA, April 16, 1975.

WINTZ, P. A., "Two-Dimensional Low Pass Filters for Images," Seminar presented at University of California, Los Alamos Research Laboratory, Los Alamos, NM, April 10, 1975.

KAK, A. C., "Noniterative Impedance Profile - Impulse Response Relationships and Signal Processing," presented at seminar on Ultrasonic Tissue Characterization, Gaithersburg, MD, May 28-30, 1975.

WORKSHOPS

HUANG, T. S., organizer and lecturer, "Image Coding, Enhancement, and Recognition," UCLA, March 3-7, 1975.

HUANG, T. S., organizer and lecturer, "Digital Techniques in Spectral Analysis, Estimation, and Filtering," a Computer Workshop, Purdue University, March 10-14, 1975.

HUANG, T. S., organizer and lecturer, "Digital Filtering - Applications to Speech and Image Processing," UCLA, May 5-9, 1975.

WINTZ, P. A., organizer and lecturer, Technology Service Corporation - Wintek Corporation, (co-sponsored), Image Processing Short Course, Zurich, Switzerland, March 17-21, 1975.

WINTZ, P. A., organizer and lecturer, Wintek Image Processing Short Course, Palo Alto, CA, April 15, 1975.

WINTZ, P. A., organizer and lecturer, Wintek Image Processing Short Course, Lafayette, IN, May 13, 1975.

BOOKS

HUANG, T. S. (ed.), Image Processing and Digital Filtering, to be published by Springer - Verlag, 1975.

KAK, A. C. and ROSENFELD, A., Digital Picture Processing, Academic Press, New York, Fall 1975. (In print)

STAFF

CO-PRINCIPAL INVESTIGATORS

P. Wintz
T. Huang

PROFESSORIAL

G. Cooper
K. Fukunaga (on leave)
D. Landgrebe
A. Kak
O. Mitchell

POSTDOCTORAL RESEARCHERS

M. Kaveh
W. Kelly

GRADUATE RESEARCHERS

S. Berger
J. Burnett
D. Chan
Wm. Chan
F. H. Chen
P. L. Chen
R. Clark
C. Myers
P. Narendra
B. O'Connor
D. Panda
A. Salahi
G. Sherman
L. Stanfield
G. Tang
E. Wiswell
M. Yoo

RESEARCH ENGINEER

W. Robey

UNDERGRADUATE RESEARCHERS

J. Besemer
C. Buckwalter
M. DeMoney
R. Johnson
P. Miller
J. Schwab
J. Uban

ELECTRONIC TECHNICIANS

D. Azpell

SECRETARIES

M. Barbour
M. Claire

MISSION
of
Rome Air Development Center

RADC is the principal AFSC organization charged with planning and executing the USAF exploratory and advanced development programs for information sciences, intelligence, command, control and communications technology, products and services oriented to the needs of the USAF. Primary RADC mission areas are communications, electromagnetic guidance and control, surveillance of ground and aerospace objects, intelligence data collection and handling, information system technology, and electronic reliability, maintainability and compatibility. RADC has mission responsibility as assigned by AFSC for demonstration and acquisition of selected subsystems and systems in the intelligence, mapping, charting, command, control and communications areas.